Universidade de Lisboa

Faculdade de Farmácia

Breast Cancer

New Screening Biomarkers and Detection Methods

Filipe Alexandre Cleto dos Santos de Sousa Rodrigues

Mestrado Integrado em Ciências Farmacêuticas

2019

Universidade de Lisboa
Faculdade de Farmácia

Breast Cancer

New Screening Biomarkers and Detection Methods

Filipe Alexandre Cleto dos Santos de Sousa Rodrigues

Monografia de Mestrado Integrado em Ciências Farmacêuticas apresentada à Universidade de Lisboa através da Faculdade de Farmácia

Orientador: Doutora Ana Cristina Ferreira da Conceição Ribeiro, Professora Auxiliar

2019

Resumo

No mundo moderno, o cancro da mama é o cancro mais diagnosticado em mulheres e o segundo tipo de cancro que mais vidas femininas colhe, e como tal para além do desenvolvimento de boas terapêuticas é igualmente, ou talvez ainda mais impactante o investimento em bons e fidedignos métodos de diagnóstico.

Vários métodos têm vindo a ser desenvolvidos, no entanto um campo de investigação em particular tem demonstrado resultados interessantes nas últimas décadas no que toca a detecção de situações de cancro: estudo do glicoma.

O glicoma tem sido um campo de investigação relativamente recente que para além de permitir compreender o papel da associação lipica/proteica com hidratos de carbono, seja a nível de interação célula-célula seja a nível célula-moléculas, tem constituído uma importante ferramenta de investigação relacionado com a prevenção e acompanhamento do desenvolvimento da situação clínica do cancro da mama. Este tipo de cancro relaciona-se com o glicoma de forma a que quando o primeiro é uma realidade o segundo se encontra alterado, resultando numa alteração da sua glicosilação chamada aberração glicómica, o qual será posteriormente detectado como forma de diagnóstico ou prognóstico, apontando então para a importância do estudo do perfil glicano.

Uma das formas de detectar o glicoma ou mais especificamente, os glicanos, envolve o recurso a lectinas que são proteínas que existem abundantemente na natureza e que possuem a habilidade de não só reconhecer os sacarídeos das primeiras estruturas mencionadas como igualmente de se ligarem a estas.

Nesta monografia será realizada uma análise sobre o que são o glicoma e as lectinas, relativamente à sua parte conceitual, à forma como interagem e acima de tudo, que vantagens trazem para a investigação de tumores, em particular para os localizados na mama. Para além disto serão também estudados métodos de idenficação adicionais que são comumente utilizados em conjunto com as lectinas de forma a facilitar e aumentar sensibilidade e eficácia da detecção de glicanos.

Palavras-chave: Cancro da Mama; Glicosilação glicómica; Aberração Glicómica; Biomarcadores; Lectinas

Abstract

In the modern world, breast cancer is the most commonly diagnosed cancer in women and the second type of cancer that most female lives reap, and as such beyond the development of good therapies is equally, or perhaps even more impactful, investment in good and reliable diagnostic methods.

Several methods have been developed, however one field of research in particular has shown interesting results in recent decades regarding cancer detection: glycome study.

Glycome has been a relatively recent field of research that, in addition to understanding the role of lipid / protein association with carbohydrates, whether at the cell-cell interaction level or at the cell-molecule level, has been an important research tool related to the prevention and monitoring of the development of the clinical situation of breast cancer. This type of cancer is related to the glycome so that when the former is a reality the latter is altered, resulting in a change in its glycosylation process called glycemic aberration, which will later be detected as a form of diagnosis or prognosis, pointing then for the importance of studying the glycan profile.

One way of detecting glycome, or more specifically, glycans, involves the use of lectins which are proteins that exist abundantly in nature and which have the ability to not only recognize saccharides of the first structures mentioned but also to bind to them.

In this monography an analysis will be made about what are the glycome and the lectins, in relation to their conceptual part, how they interact and, above all, what advantages these structures bring to the investigation of tumors, particularly those located in the breast. In addition to this, additional identification methods that are commonly used in conjunction with lectins will also be studied in order to facilitate and increase sensitivity and effectiveness of glycan detection.

Keywords: Breast Cancer; Glycome Glycosylation; Aberrant Glycome; Biomarkers; Lectins

Acknowledgments

With the delivery of this monography I begin the conclusion of the final stage of this long journey which was the Integrated Master in Pharmaceutical Sciences and as such, thanks will have to be made for all the help and support given.

Firstly, I would like to thank my advisor, Prof. Ana Cristina Ferreira da Conceição Ribeiro for the great patience she had with me and for always being available to help me, and whose guidance and bibliographical indications were fundamental for the preparation of this monograph.

I would also like to thank my parents, Isabel and José, for always being so worried about me, for offering constant help and support, and especially for helping me to stay calm when everything was more complicated. To my brother Diogo, who over the years has always been by my side in the worst and best situations, always giving me his unconditional support always with a great humor, revealing himself as a great friend.

To Farmácia Holon Covilhã and its team for betting on my training as a future pharmacist, always available to answer any questions.

To my longtime friends Rui Ferreira and Pedro Fernandes, for sharing so much with me and for always extending a helping hand to me in the highest times I needed, regardless of the inconvenience that might cause them.

To my equally friends, Pedro Neves and Rui Franco for making me have such a good time and always providing distractions when needed and especially listening to me whenever I needed them.

To Luis and Daniela for always inviting me to go out with them, for providing such pleasant and enjoyable table top game nights and especially for being such good friends.

To my friends I take from college, especially Marta Anselmo, for always pulling me up when things seemed the hardest and for being such a good confidant; Joana Luís for being such a good friend always ready for helping me in any way she could. Inês Santos, for being the first friend I made and for supporting me and showing positivity anytime I needed; Iulia Jucov, who helped me a lot to get to where I find myself by pulling me as much in my studies, as always being such a fun friend; Dany, Chico, Bruno and Zé who, even knowing less time, have helped me to relax in the moments of greatest tension; Margarida Oliveira for having endured me so often in group work.

To my godmother Cátia Henriques for being always there for me, helping with anything I needed and guaranteeing that I had a pleasant time where and whenever we were together.

To Ana, Silvia, Rita and Raquel for being a phenomenal group of friends, with fantastic humor that made me laugh a lot and also for the huge support they have given me lately.

To Mariana for being the fantastic and dear person she is and for being a pleasant surprise I didn't expect, helping to make this final journey more fun and light-weighted, at a time when stress prevailed, not forgetting the enormous support she has given me.

Finally, I would like to dedicate my monography to all these people as well as my grandparents Margarida, Delfim, Fernanda, Celso, my great grandparents Albertina and Aníbal and my cousins Álvaro, Alexandra, João, Mariana, Guilherme, Paulo, Marlene, António, Afonso, José Fernando, José and Anita, as well as my godparents José Alexandre and Cândida Rueff.

Abbreviations

DCIS – Ductal carcinoma *in situ*

LCIS – Lobular carcinoma *in situ*

HR – Hormone (estrogen or progesterone) Receptor

HER2 – Human epidermal growth factor receptor 2

MALDI – Matrix-assisted laser desorption/ionization

ESI – Electrospray ionization

HPLC – High performance liquid chromatography

TOF – Modem time of flight

ICR – Ion cyclotron resonance

PNGase F – Peptide N- Glicosidase F

NaBH4 – Sodium Borohydride

GalNAc – α-N acetyl-d-galactosamina

UV – Ultra-violet

SAMs – Self-assembled monolayers of thiol-glycans

DNA – Deoxyribonucleic acid

SPR – Surface plasmon resonance

PDMS – Polymethylsiloxane

Asn – Asparagin

GlcNAc – N-acetylglucosamina

Ser – Serine

Thr – Threonine

Man – Mannose

ER – Endoplasmatic Reticle

UDP – Uridine diphosphate

PTS – Proline, threonine and serine

ppGaNTases – UDP-N-acetylgalactosamine:polypeptide N-acetylgalactosaminyltransferases

FDA – Food and Drug Administration

CA15-3 – Cancer antigen 15-3

MUC1 – Mucin 1

CA27-29 – Cancer antigen 27-29

CEA – Carcinoembryonic antigen

SEA – Sperm protein, enterokinase and agrin

VNTR – Variable number tandem repeat

ECD – Extracellular domain

TMD – Transmembrane domain

CT – Cytoplasmatic tail

TA-MUC1 – Tumor-associated Mucin 1

IgA1 – Immunoglobulin A1

CRD – Carbohydrate recognition domain

ConA – Concanavalina A

ELISA – Enzyme-linked immunosorbent assay

HPA – *Helix pomatia* agglutinin

SNA – *Sambucus nigra*

MAL-II – *Maackia amurensis* lectin II

ELLA – Enzyme-linked lectin assay

Index

List of Tables

List of Figures

1. Breast Cancer Introduction

1.1. Etiology

Breast cancer is a major public health problem, since it is one of the most frequent forms of cancer that affect women all over the world. This kind of malignancy has the particularity of showing divergent pathological characteristics: there are situations where the prognosis is quite positive, while in another cases the tumor presents itself with a very fast growth. (1)

The formation of these kinds of tumors starts usually with an augmented proliferation from the ductal cells which evolve to benign tumors or metastatic carcinomas, if there is enough carcinogenic stimulation. (2) Of the many risk factors that lead to the outcome before mentioned the most worth mentioning are: long time fertility (consequence of early menarche ages and menopause in old ages), the use of preventive pregnancy hormones or hormone replacement therapy, obesity after menopause, alcohol consumption and physical inactivity. On the other side of the spectrum, having children and breast-feeding have a preventive role in breast tumor development. (3)

Breast cancer can be divided into 2 types: *in situ* and invasive cancer. The in situ cancer can also be splitted in two categories: ductal carcinoma *in situ* (DCIS) and lobular carcinoma *in situ* (LCIS) or lobular neoplasia. Some cases can also include a combination of both carcinomas or an undisclosed origin. The ductal carcinoma is one of the most frequent cases (around 80% of the cases *in situ* diagnosed by the American Cancer Society between 2010-2014) and applies to situations where normal functioning breast duct cells change into abnormal ones, which can lead to duct and lobule expansion, even though DCIS not always becomes invasive because it grows slowly. The lobular *in situ* carcinoma, LCIS, is more uncommon and it's characterized by abnormal cells growing and expanding the breast lobules, resulting in an increased risk factor for developing invasive cancer, even though LCIS not being a precursor of that situation. (4)

Regarding the invasive cancer, this is one of the most frequent types of cancer and although is commonly seen as a single disease, actually there are 21 histological subtypes and 4 molecular subtypes (Table 1). These last subtypes have been identified and studied by using gene expression profiling techniques, but because of the great complexity and expensiveness of this procedure, the identification has been usually made by using biological markers such as the presence or not of estrogen or progesterone receptors (HR+/HR-), excess levels of

HER2 (human epidermal growth factor receptor 2) and additional copies of HER2 gene (HER2+/HER2-). (4)

Table 1 – Invasive Cancer Molecular Subtypes

Subtypes	Characteristics	References
Luminal A (HR+/HER2-)	• Most frequent subtype (71%); • Less aggressive and slow growth; • Best prognosis; • Most responsive to anti-hormone therapy.	• (5)
Luminal B (HR+/HER2+)	• Much less frequent subtype (12%); • Extremely positive for Ki67 (indicates a great proportion of active cell division) or HER2; • Worst prognosis than Luminal A.	• (5)
Triple negative (HR-/HER2-)	• Much less frequent subtype (12%); • More common in black, premenopausal and BRCA1 gene mutated women; • Worst prognosis of the 4 subtypes (no targeted therapies).	• (5)
HER2-enriched (HR-/HER2+)	• Less frequent subtype (5%); • Most aggressive and faster growth; • Bad prognosis compared to other HR+ subtypes (best results with targeted therapies)	• (5)

Adapted from Atlanta: American Cancer Society, 2017 (4)

Even though this type of cancer has been studied for a long time, there are two theories regarding the initiation and progression of breast cancer, being both of them supported by data, however neither of them can explain on their own and completely the beginning of this situation. The first one is *the cancer stem cell theory* which suggests that all tumor subtypes derive from the same stem cells or progenitor cells which will differentiate into diverse tumor phenotypes as result of different genetic and epigenetic mutations. The second theory is the *stochastic theory.* The stochastic theory explains that each tumor subtype starts either from a stem cell, a progenitor cell or even a differentiated cell. Then the accumulation of random mutations in the breast cells will result in the formation/transformation into tumor cells (when enough mutations have happened). (2)

1.2. Epidemiology

Breast cancer has had a huge impact in women around the world, and citing the American Institute for Cancer Research, it's the fifth most common cause of death, cancer-related, adding around 1.7 million new cases (25% of the new cancer cases detected in women) just in 2012. (6)

According to the *Ghoncheh et col.*, the developed countries (specially Northern America) have a bigger incidence in terms of breast cancer since next to 50% of patient cases

and specifically 38% of deaths happened on those countries in 2015. Nonetheless, South America, Asia and Africa have had increasing numbers of this kind of cancer in the last four years, but more than 50% of the eligible women have been screened. The problem though, involves women that are immigrant or with low economic power, which have not been screened in an adequate rate, resulting in an association between lesser survival rate and lesser privileged areas and/or people. (3) And so, the survival rate in a 5 year scale will be bigger in countries (e.g. USA, Australia and Canada) that have a bigger breast cancer incidence but that are more developed, while countries in development with lower incidence and/or less medical care to support regular screenings and consequently, treatment (e.g. Slovenia and Singapure) will have also less survivors/bigger mortality rates, like it was referred previously. (1)

2. Breast Cancer Carcinogenesis

2.1. Glycome as an Overview

Every cell has a cytoplasmatic membrane in which exterior is set a series of glycolipids and glycoproteins complexly organized with oligosaccharidic moieties – the **exoglycome** (or just **glycome**). These oligosaccharides can be seen as a code that is involved in a major form in cell-cell and cell-molecule interactions. (7) Even though the glycome is considered an analog of the proteome and genome, the first of three above mentioned, it's much more complex because of the immense variety of glycans which makes not only the uncover of glycome impossible by only using proteome but also for the connectivity between glycome and genome, proteome and metabolics (resulting in a difficulty in the classification). (8)

The chains of saccharides presented in **glycome** can be referred as **glycans** and, like it was mentioned before, they have an extremely important role being considered one of the four building pillars of the cell composition, in conjunction with nucleic acids, proteins and lipids. **Glycans** have a wide variety of functions starting from assisting protein trafficking and folding, cell adhesion, immunity system modulation, pathway signaling (working like a postal code for route and deliver) and also can act as protective layer of the outside cell or intervene in the infectivity process of bacteria and viruses. (9)

A glycan can be found in every mammalian cell or even on body fluid and it is formed by a series of monosaccharides (which can have different forms with the same composition – stereo and regioisomers (10)) linked to each other through glycosidic bonds, but not

exclusively, since covalent bonds can be made with proteins and lipids, forming in this last case **glycoconjugates**, which have multiple classes (Figure 1), being that *O*- glycans (serine/threonine residue connection) and *N*-glycans (asparagine residues connection) are the most predominant post-translational modifications of proteins. Each copy of glycoprotein with a specific glycan is defined as **glycoform**. (9)

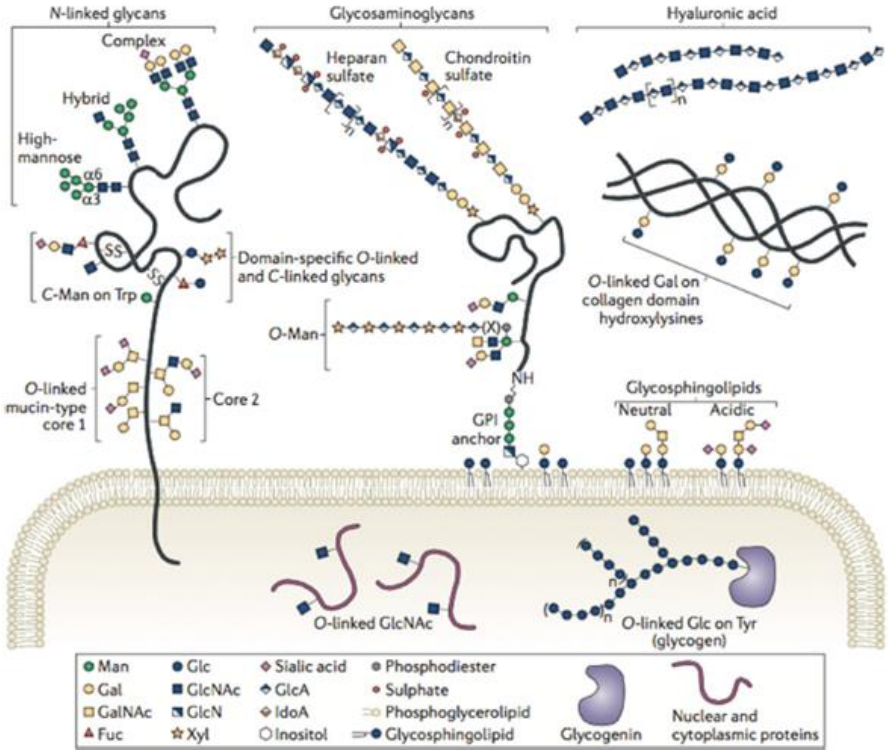

Figure 1 – Schematic representation of the mammalian glycome. Different glycoconjugates (glycans appended to proteins and lipids). There can be found in a great number O-linked GlcNAc on either cytoplasmic or nuclear proteins, but the most common post-translational modifications of extracellular proteins are N- and O- glycans. According to glycobiology nomenclature, the sugars are graphically differentiated through different geometric shapes and colors.
Abbreviations in the figure: GalNAc - N-acetylgalactosamine; GlcNAc - N-acetylglucosamine; Gal - galactose; Glc - glucose; Man - mannose; Fuc - fucose; Xyl - xylose; GlcA - glucuronic acid; IdoA - iduronic acid; GlcN - glucosamine.
Adapted from Taron and Rudd, 2016 (9)

Another particularity of the glycan refers to the complexity of its own constitution since the assembly of the structure follows elaborate biosynthetic pathways and also because of the capability of changing in a fast manner its original form in response to environmental stimuli, illustrating in this way the interchangeability and variability of the glycan without even affecting the genome. The genome is related to the glycome in a way that it encodes proteins involved on formation and recognition of glycans, and enzymes (glycosyltransferases) which intervene in glycosylation. (9) As it was mentioned at the beginning of the present paragraph, in order to assemble the wide variety of glycan structures

there are metabolic pathways that must be followed which must include: the formation of precursors from nucleotide sugars (actived monossaccharides that will intervene in glycosylation reactions by transfering glycosyl groups), sugar transporters (guarantee the existence of enough cell intermediates to form the precursors), and glycosyltransferases (selective expression affects what glycan sequence is present), responsible to bring the saccharides to the target protein leading to the formation of the pretended glycan structure, as well as many other proteins involved in producing a glycoprotein adequate for a certain purpose. These pathways work in a manner that permit, not only additional manipulations on the final product but also catabolic ones, which can be associated to normal turnover, degradation or specific changes culminating on active glycoforms, with salvage pathways as the feeding mechanism of the glycoprotein metabolism. (11)

In order to achieve progressively a bigger understanding about glycomics (study of glycans existent in all biological systems), it was necessary to do what has been done before with the genome, which was proceed to develop tools that allow profiling and consequently, analysis of the glycome. Several tools field have been used since the beginning of glycomics such as: **mass spectrometry**, **chromatography** and **glycan and lectin array**, which will be described in "Analyzed Methodology".

2.2. N- and O- Glycosylation

As it was early and briefly mentioned at the introductory section of glycans, glycosylation is a step by step process in which are formed complex and diversified structures through sequential attachment of saccharides, constituting the most frequent post-translational modification of both proteins and lipids. (12) Glycosylation is a very frequent process through which proteins and lipids often go (around 70%) and focuses on the cells' surfaces (the situation intended to study in this article) and in extracellular matrices, revealing that glycan production has a more serious and significant impact in disease situations than protein synthesis (8), so if glycosylation patterns are changed, this could represent a major change in the process resulting in problems like carcinogenesis. (12)

Glycoproteins exhibit different kind of glycan binding, the N- and O-glycans, with a lesser prevalence in C-glycans. The N-glycans are covalently linked to the protein component at asparagine (Asn) which is consequently attached to GlcNAc (N-acetylglucosamine) while

O-glycans have serine (Ser) or threonine (Thr) linked to GalNac (α-*N*acetyl-d-galactosamine). (11)

2.2.1. N-Glycosylation

Even though N-glycans are not identical they share a common core constituted of Manα1,6(Manα1,3)Manβ1,4GlcNAcβ1,4GlcNAcβ1-Asn-X-Ser/Thr (with the generation of two antennae from the core) which extends to three different categories: **oligomannose**, **complex** and **hybrid**. In **oligomannose** N-glycans the residues extended from the core are only composed by mannose (13), while in **complex** glycans, the antennae terminations are composed of sialylated N-acetyllactosamine trisaccharide with fucose attached to Glc*N*Ac. (11) Finally, **hybrid** N-glycans have only mannose extension on Manα1,6 arm and on Manα1,3 arm one or two extensions from the core like in **complex** types (Figure 4). (13)

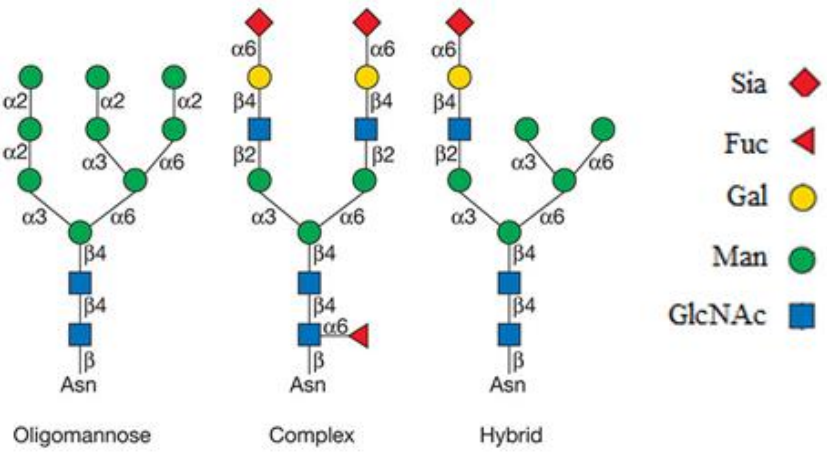

Figure 2 – General Structure of the 3 types of N-Glycans. N-Glycans can be of three principal types: oligomannose, complex and hybrid. Every single N-Glycan has a common core of Man3GlcNAc2Asn and complex N-Glycans can have no more than six branches (initiated by GlcNAc and elongated with LacNAc). Abbreviations in the figure: Sia – Sialic Acid; Fuc - Fucose; Gal - Galactose; Man – Mannose; GlcNAc – N-acetylglucosamine; Asn - Asparagine
Adapted from Varki *et al*., 2017 (13)

In the *N*-glycosylation, the synthesis can be divided into two main steps which happens in two different organels: Endoplasmatic Reticle (ER) and Golgi Complex. The first step: formation of the glycan precursor occurs on the ER membrane (cytosolic side), where the enzymes involved use a lipid carrier of dolichol pyrophosphate, following then a series of trimming and processing steps. Then a precursor oligosaccharide preformed (Man5GlcNAc2 results from the action of three glycosyltransferases, first the intervention of Glc*N*Ac

transferase then by mannosyltransferases) is assembled on the lipid carrier (two N-acetylglucosamine, one phosphate, and five mannose residues from UDP sugars are sequentially and in a not simultaneous way added to the dolichol pyrophosphate). Soon after, tunicamycin will hinder the formation of the N-linked saccharides with the follow up of a flipping of the dolichol pyrophosphoryl oligosaccharide by a flippase putting it on the luminal side. Now in the lumen, it's added an additional four mannose and three glucose residues obtaining the complete formation of Glc3Man9GlcNac2 precursor which will be latter transferred to specific Asn residues in the target Asn-X-Ser/Thr sequence of both secretory and membrane proteins. It is noteworthy to mention that this glycosylation does not happen in all Asn-X-Ser/Thr sequences, since the folding of a segment of a protein containing the previously mentioned sequence can be enough to prevent Glc3Man9GlcNAc2 transference. (11,13–15)

Regarding on the second phase of N-glycosylation, it will involve processing by glycosidases and glycosyltransferases (which will use activated sugars as substrates) starting where the first phase left behind, ER lumen, and continues in the Golgi complex. So nextly, after it was added the oligosaccharide to the protein, three glucose residues plus one mannose residue will be removed sequentially (first one glucose residue, then two and finally one mannose residue), following a re-insertion of a glucose by a glycosyltransferase in either unfolded or misfolded proteins. In order to prevent a recurrent folding in these proteins, ER have two lectins (calnexin and luminal calreticulin) which will bind specifically to them. From this point two things can happen: the proteins can stay linked to the lectins or they can separate themselves and become deglucosylated again. In the latter case, if the proteins evolve in a matter that their folding is done properly, they will not pass for the same process previously mentioned (folding prevention mechanisms) but instead they advance to next step, Golgi complex mobilization. When these proteins enter the Golgi they have at least one Man8GlcNAc2 chain and can suffer different modifications, while go through this organelle to arrive the outer cell, because of the diversity of enzymes that exist on *cis* (mannosidase I), *medial* (GlcNAc transferases) and *trans* (galactosyltransferase and sialyltransferase) cisternae. In the end, several variations on N-linked oligosaccharides structure can result from the crossing of ER and Golgi complex since they suffer modifications of different enzymes (specially on Golgi) which stop only when the enzymes can longer access to the areas likely to be manipulated, finishing with these N-glycan being secreted or embedded in plasma membrane. One very important factor that also affects the glycosylation (besides exposition to

different enzymes) relates to the type of cell it is, its physiological state and which glycosylation genes are being expressed. (13,15)

2.2.2. O-Glycosylation

The O-glycosylation characteristic of mucins (glycoproteins produced by glandular epithelia which are confined to the membrane or secreted by the cells (16)) is a modification that has a major role regarding protein processing, secretion, stability and function, so any problem in this process can affect negatively its normal functioning resulting in aberrant behavior, resulting and diseases which is very common in adenocarcinomas (more fully explained at Aberrant Glycosylation section). (17)

Just like the N-glycan, so do O-glycans are a common target of glysosylation (second preferred after N-glycans) and can be divided into eight categories, taking in account the different core structure existing. These kind of glycosylation is also known as mucin type glycosylation since it is the biggest modification regarding saccharides addition happening in the all group, but it is worth mentioning that there are other types of O-glycans which are not mucins, even though they are not going to be addressed in this article. The mucin type O-glycans are positioned in a variable number of repeated domains (called PTS domains because of its constitution: proline, threonine and serine (17)), which, depending of the concerned mucin, can have a different size and sequence and unlike N-glycans, O-glycans do not have a known sequence for peptide recognition. (11)

Contrary to what happens in N-glycosylation, the process where modification of the aminoacids (serine and threonine) occurs only on the Golgi complex, more specifically at *cis* and *trans* cisternae and it is initiated by peptidyl GalNAc-transferases. After glycosylation starts, this process will follow two rules: the sequence order of supposed to be glycosylated sites and epigenetic regulatory mechanisms mediated through enzymatic competition. But, the way O-glycosylation happens in each cell is determined by the correspondent set of cellular set of glycosyltransferases, by their sugar donor and acceptor particularities, their sequential action and location on the Golgi complex. (18)

In mucin-type O-glycosylation, as it was mentioned in the paragraph above, the modification starts with action of a large family of glycosyltransferases named UDP-N-acetylgalactosamine:polypeptide N-acetylgalactosaminyltransferases (ppGaNTases, EC 2.4.1.41) which will provide GalNAc sugar from the sugar donor UDP-GalNac to Ser or Thr

residues at the level of the hydroxyl group, resulting in the formation of Tn Antigen. This happens in specific sites with intervention of several ppGaNTases and depends on both the structure and position of O-glycans that were formed before. In terms of the previously referred family, every member are type II transmembrane proteins is capable of adding GalNac in this type of O-glycosylation, being that there are 20 known enzymes in humans, 19 in mice and 12 in *Drosophila* flies. The complexity associated to this family is related to the fact that each member has specific expression patterns, but even though some are present in many developing tissues, others have a very limited and specific expression, space and time wise, revealing also preference for certain kinds of substrates and target sites in those proteins for GalNAc addition. (17,19)

Following the GalNAc addition, the next step is addition of saccharides, sequentially, in order to extend the sugar chain, being that Core 1 or T antigen structure is the most common extension, catalyzed by T-synthase/C1GalT1 (core 1 β1,3-galactosyltransferase) responsible for transferring galactose monosaccharides to GalNAc via β1,3-linkage. In order to C1GaT1 be active (in mammals) and perform their function, a chaperone from ER, Cosmc, is needed, and if there is any flaw in this enzyme the synthesis of core 1 is affected, but only this sugar chain. (17)

Another extension possible to form is core 3 through β1,3-N-acetylglucosaminyltransferase 6 (expressed most predominantly in the digestive tract), responsible for forming a β1,3-linkage between the first added GalNAc and GlcNAc. Both this structure and core 1 (principle extension structures formed in O-glycosylation) can then be modified with the addition of GlcNAc, via β1,6-*N*-acetylglucosaminyltransferases, forming core 4 and core 2, respectively. In the mammals, there are 3 types of this last mentioned enzyme, where two can catalyze the formation of core 2 but only one can catalyze the formation of both core 2 and core 4 structures. (17)

In addition to the above mentioned structures and respective modifications, it is possible to generate longer linear or branched structures with the addition of more galactose and GlcNAc like it has happened before, or through fucosylation or sialylation (usually how the structure ends. This all process can be summarized in Figure 6. (17)

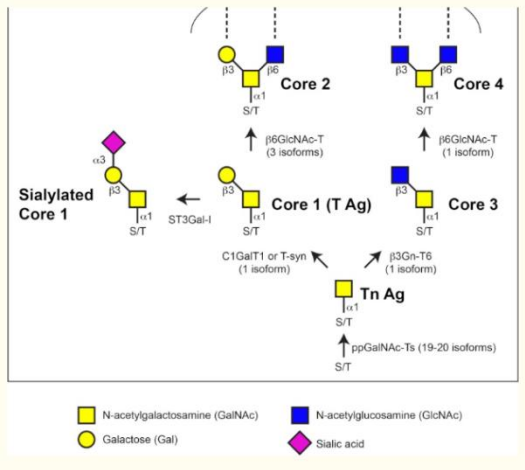

Figure 3 - Biosynthesis of mucin-type O-glycans. The O-glycosylation in mucins is initiated with the addition of GalNAc to the hydroxyl groups of serine or threonine in protein substrates, forming the Tn antigen (Tn Ag). After this it follows the core structures formation through sugar addition. Enzymes responsible for the synthesis of the Tn antigen, core 1 (T antigen (T Ag)), core 2, core 3, core 4, and sialylated core 1 structures are represented in the figure. The numbers in parentheses represents the mammals' isoforms.
Adapted from Tran and Hagen, 2013 (17)

Although there is very information about this process, there are some problems associated to its analysis, regarding the methods to do it: the fact that the enzyme family responsible for the GalNAc sugar addition is so large in size and also because of the consequent functional redundancy subjacent related to the formation of O-glycans, it's hard to analyze this process, especially if resorting to single gene knock-outs. Besides this point, the difficulty of studying this kind of protein modification is incremented by not knowing the sequence in which will be insert GalNAc and for the fact that there is no method for detecting all O-glycans by just using one single reagent neither an enzyme capable of removing all these glycans. (17)

2.3. Aberrant Glycosylation and Biomarkers

2.3.1. Biomarkers

After the brief explanation and addressing of the normal course of glycosylation it is necessary to go deep and approach the aberrant situations where glycosylation, either N or O focused, don't work in a correct manner. These situations are referred as aberrant glycosylation which are very present in disease cases, namely in breast cancer cases. (20) As been said previously, glycosylation differs in a great way, comparatively to protein or nucleic acid synthesis, since the first one, contrary to the others, doesn't follow a template, and

besides that, the complexity consequent extends widely the impact that oligosaccharides have in cell-cell and cell-matrix interactions regarding cancer. This factor allied to the fact that human serum proteome is hugely composed of glycoproteins will be very important to achieve screening biomarkers and detection methods of those in breast cancer, since proteins are secreted or leaked from tissues or blood cells to the circulatory system, which grants the possibility of being later analyzed through serum collection (easy to harvest from peripheral blood with minimal risk to patient), allowing to study if and how their correspondent oligosaccharides structures were affected in case of stimuli resultant of pathophysiological changes, in terms of breast cancer. (21) Aberrant glycosylation is not unique to breast cancer but are present in all type of cancer, which makes even more important the study of glycome and its fluctuations of structures, especially since a great number of glycosyl epitopes can be considered tumor antigens.(22)

In this work, it will be related breast cancer biomarkers and glycosylation modification, explaining how those events affect the biomarkers mentioned in order to make them effective or promising tools to detect anomalies. In the lectins context, it will be explained how those biomarkers are detect through use of the above referred proteins and other complementary methods to increase reliability in screening tests.

Regarding biomarkers for breast cancer screening, thanks to the development of molecular biology in terms of tools and methods, there has been the discovery of many biomarkers, which have a huge importance since they can to used (with other strategies), in a very minimal invasive way, to establish prognostics and help predict the evolution of the health condition of the patients. (23) The biomarkers that have been used currently clinically are characterized to be most accurate in situations of widespread cancer because usually in early stages it's very hard to detect most of them, which are more expressed in benign situations. Another particularity of biomarkers contemplates the fact that biomarkers with high specificity and sensitivity are hard to identify in most situations as a consequence of tumor molecular heterogeneity and diversity of tumor stages in the same tissue/organ, even though this is something that has been studied for many years and quite intensively. Nevertheless, even though biomarkers have limitations in their usage a few have an acceptable specificity and sensitivity and so, are approved to be used clinically by FDA (Food and Drug Administration), which are represented in Table 2. (21)

Table 2 – List of FDA-approved breast cancer biomarkers currently used in clinical practice

Biomarker	Full Name	Clinical Application	Year of FDA Approval
CA15-3 (MUC1)	Cancer antigen 15-3	Monitoring Therapy	1997
CA27-29	Cancer antigen 27-29	Monitoring Therapy	2002
CEA	Carcinoembryonic antigen	Monitoring therapy, detecting recurrence	1985
HER2/neu	Human epidermal growth factor receptor 2	Therapy choice	1998

Adapted from Kirwan, 2015 (21)

Of the biomarkers mentioned in Table 2, this article will focus on: CA15-3. Besides this biomarker another one that has been studied in the last year will be approached for its promising results in breast cancer screening, allowing a glimpse into the state of the art.

2.3.1.1. As Glycoproteins

2.3.1.1.1. CA15-3 (MUC1)

Cancer antigen 15-3 or MUC1 (episialin, and milk mucin antigen are also synonyms for this biomarker) is a transmembrane mucin (high molecular mass glycoprotein) which is present in the majority of glandular or luminal epithelial cells (mammary gland, esophagus, stomach, duodenum, pancreas, uterus, prostate, and lungs), exhibiting a protective role regarding the underlying epithelia, anti-adhesive properties and grants a physical barrier limiting pathogenic action, thanks to its glycosylated branches (negatively charged). MUC1 has intensive oligosaccharide addition via O-glycosylation or can be moderately N-glycosylated, and through the oligomerization of the carbohydrate chains it is achieved a gel that grants lubrication and protection, avoids desiccation, pH changes and intervention of pathogens, like microbes, regarding the underlying epithelia. (21,24) CA15-3 was first identified in human milk, being shed from lactating mammary epithelial cells but can also be found with increased levels in serum of breast cancer patients. (21)

Regarding the structure of this mucin, it is composed by a single polypeptide chain, which can be divided into 2 regions (Figure 8): the N-terminal subunit also known as MUC1-N (longer subunit) and the C-terminal subunit or MUC1–C (shortest subunit). (24)

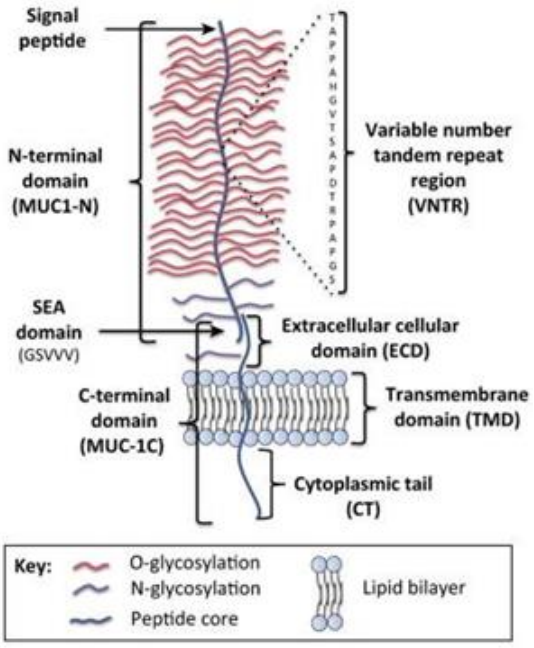

Figure 4 – Schematic structure of a full length MUC1. The two terminal subunits (MUC1-N and MUC1- C) form a stable heterodimeric complex, through association with the SEA domain. MUC1-N encompasses the signal peptide, the VNTR (Variable Number Tandem Repeat, which has 20 amino acids that are going to be extremely O-glycosylated at both serine and threonine residues - red) and the SEA domain. MUC1-C includes the extracellular domains (ECD), the cytoplasmatic tail (CT) and the transmembrane domain (TMD). Both terminal subunits, MUC1-C and MUC1-N, are moderately N-glycosylated at asparagine residues – violet. Abbreviations in the figure: MUC1- Mucin 1; SEA – Sea urchin sperm protein, enterokinase and agrin.
Adapted from Nath and Mukherjee, 2014 (24) and Carson, 2008 (25)

The longer subunit is constituted by 2 parts: the SEA domain (Sperm protein, Enterokinase and Agrin) where, soon after translation, happens autoproteolytical cleavage (at GSVV motif) because of conformational stress, leading to the separation of the two subunits firstly mentioned; the VNTR region (variable number tandem repeat). Nevertheless both MUC1-N and MUC1-C stay after associated extracellularly thanks to hydrogen bonds. The last mentioned part of MUC1-N, the VNTR region, has 20 amino acids where it will occur O-glycosylation in great extension.(24)

The shortest subunit is constituted by 3 parts: ECD (extracellular domain), TMD (transmembrane domain) and CT (cytoplasmatic tail). The ECD has a region where N-glycosylation will happen, even though this type of glycosylation also has a region where occurs in MUC1-N. (24)

Regarding aberrant glycosylation in this situation, there are two types of MUC1: the normally expressed one and the tumor associated MUC1 (TA-MUC1). The big difference between those two is that the first one has normal breast tissue producing O-glycans core 2-based, while TA-MUC1 contains Core 1 O-glycans as consequence of the loss of Core 2 β6-GlcNAc-transferase activity (Figure 9).

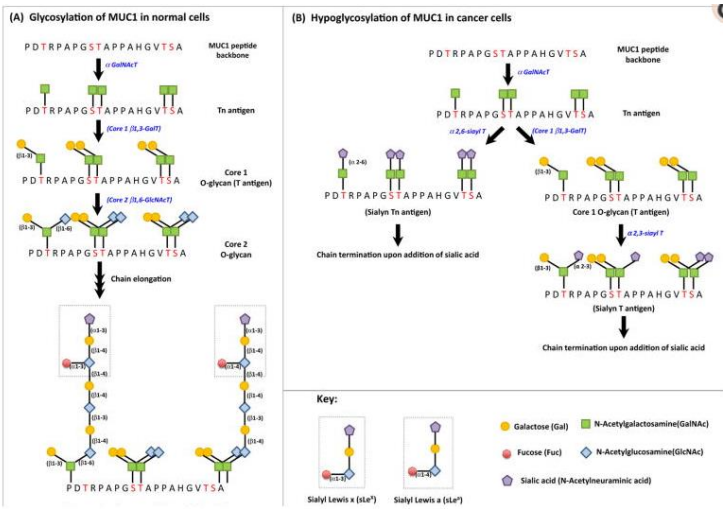

Figure 5 – Schematic of the differences between the glycosylation patterns in normal and tumor-associated MUC1. In (A) there is a representation of MUC1 glycosylation in normal cells, where GalNAc is added to serine and threonine residues, via αGalNAc transferase, leading to the formation of Tn Antigen, which is followed by galactose addition to Tn antigen, through Core 1 β 1,3 – galactose transferase forming T antigen. After this step, T antigen gets added GlcNAc forming Core 2 glycan, whose sugar branches will suffer elongation and finally, termination by addition of fucose or sialic acid to the terminal sugar. Regarding (B), the figure represents hypoglycosylation of MUC1 in cancer cells, by premature termination of elongation of the sugar branches, because of sialylation of Tn and T antigens. Abbreviations in the figure: GalNAc T – N-acetylgalactosamine transferase; GlcNAC T – N-acetyl glucosamine transferase; Gal T – Galactose transferase; Sialyl T – Sialyl Transferase
Adapted from Nath and Mukherjee, 2014 (24)

Furthermore, TA-MUC1, unlike the normally expressed glycoprotein, is extremely sialylated (could be because of augmented expression in the α2,3-transferases in breast cancer cells) which leads to termination of chain elongation and formation of truncated sugar branches. (24,26)

CA15-3 in breast cancer will be not only glycosily aberrant but also overly expressed. One important note that must be considered is that the diagnosis of metastatic breast cancer with this mucin should be done in patients with non-evaluable metastases (preferably bone dominant diseases), since the levels of MUC1 can increase as a result of necrosis and apoptosis induced by chemotherapy. One of the problems regarding the usage of MUC1 in screening is that it is not possible to predict if the increasing levels of MUC1 points to cancer evolution or if decreasing numbers of this mucin reflects effectiveness in the treatment. (26)

2.3.1.1.2. IgA1

IgA1 or Immunoglobulin A1 is an antibody that grants protection against external pathogens and antigens encountered at mucosal sites (27) and it is composed by 4 chains: 2 heavy ones and 2 light chains connected via disulphidic bridges (Figure 6). The first type of chains have 3 constant domains (Cα) and another variable domain (V), while the light chains one of both of those domains (C and V). Between the constant domains Cα1 e Cα2 in the heavy chains, there is a hinge region with an amino acid sequence with nine potential regions likely to suffer O-glycosylation (only 3 to 5 actually are glycosylated: Thr228, Ser230, Ser232, Thr225 and Thr236 – the last two less frequently), where, even though the composition may vary, the most common O-glycans are Core 1 with sialylation at the end. The IgA1 can also be N-glycosylated in 2 regions (Asn263/Asn459) per heavy chain. (28)

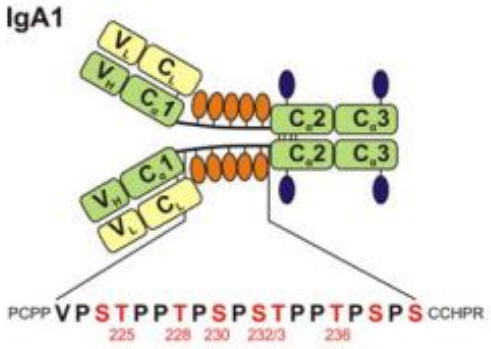

Figure 6 - Schematic of molecular structure of human IgA1. IgA1 has two light chains (L) and two heavy chains (H). There is an hinge region where between Cα1 and Cα2 (constant domains of heavy chain) are attached O-Glycans (orange oval symbols). The hinge region's sequence has two octapeptide repeats with no more than six O-Glycans attached (numbered S/T residues). Also in the heavy chain, there are two N-Glycans (blue oval symbols). Constant and variable domains of light chains (CL, VL) are yellow and heavy chains (Cα1-3, VH) are green. POGS (potential O-glycosylation sites) are marked in red, with six residues glycosylated on circulatory IgA1 being numbered.

Adapted from Stuchlová *et al.*, 2013 (28)

Even though, it is known that IgA1 suffers glycosylation, its physiological importance is still unclear besides IgA clearance by hepatocytes via the asialoglycoprotein receptor and glycan-dependent immune exclusion by secretory IgA. (27) This immunoglobulin exist in serum (as well as in tissues and mucosal secretions like breast milk) mostly as a monomer but it can also associate itself with other/s forming dimers or even oligomers (collectively polymeric IgA1). (28)

Usually aberrant glycosylation in IgA1 glycans include diminished galactose in *O*-glycans, accompanied with oversialylation or undersialylation, with the possibility of aberration of N-glycans. (27) But regarding breast cancer, the aberration according to some studies involves increased levels of terminal GalNAc (marker of poor prognosis and aggressive breast cancer case), increased bifucosylation and noticeable aberrant glycosylation in N-glycans regarding increased sialylation (analyzed from serum). (29)

In this particular case, IgA1 it is not still an approved biomarker to be used in breast cancer (usually it is more applied in nephropathy), being in a phase were it has been study if it is reliable in order to be used more frequently. For that purpose, this segment takes in account, especially, the article "Serum IgA1 shows increased levels of α 2,6-linked sialic acid in breast cancer" (and its results) which analyzes serum IgA1 glycosylation in breast cancer, in order to measure its potential as biomarker for prognostication. Looking to the results, obtained IgA1 shows some potential to be a biomarker for breast cancer detection (because it was detected altered glycosylation in IgA1 in breast cancer cases), but leaves some questions open since it is unclear if the contemplated increase in sialylation of IgA1 in breast cancer is of functional significance and also if the increased serum IgA1 levels are caused from tumor infiltrating or circulating B-lymphocytes or from residual disease. (29)

2.3.1.2. microRNA

MicroRNA or miRNA are single-stranded RNA sequences with a short number of nucleotides (between 19 to 23 nucleotides) obtained from 70 nucleotides precursors that affect and regulate, via miRNA pathway, how gene expression occurs in several physiological processes. It is also worth mentioning that even one miRNA has the possibility to target hundreds of mRNA and through that, intervene on mRNA repression by paring to complementary sequences which leads to transcript destabilization and/or translational repression. (30,31) The miRNA is encoded on the genome, where have been discovered around 42500 sequences, according to the miRBase database. (32,33)

In regards to miRNA genesis and maturation multiple steps must occur before reaching a mature miRNA. Firstly, in the nucleus, a primary miRNA (pri-miRNA) suffers transcription by the action of RNA polymerase II or III, for then being cleaved by Drosha, a class 2 ribonuclease, so it can be formed a precursor of miRNA (pre-miRNA) (30,34). After the end of the last step, the newly existent pre-miRNA is transported to outside the nucleus by the

exportin-5 and loaded onto Dicer so the loop can be cleaved forming a double-stranded structure with miRNA and antisense miRNA. The antisense miRNA in most cases is degraded, remaining just the mature miRNA strand which will be added to the miRNA-induced silencing complex (mRISC). This action will lead to gene silencing through mRNA cleavage or translational repression (Figure 7), as it was mentioned before, or even translational induction. In order to regulate the mature miRNA levels, there is binding to these structures by circular RNA, pseudogenes, and lncRNAs (class of RNA molecules with no more than 200 nucleotides), which wil prevent constant miRNA binding to target mRNA. (30)

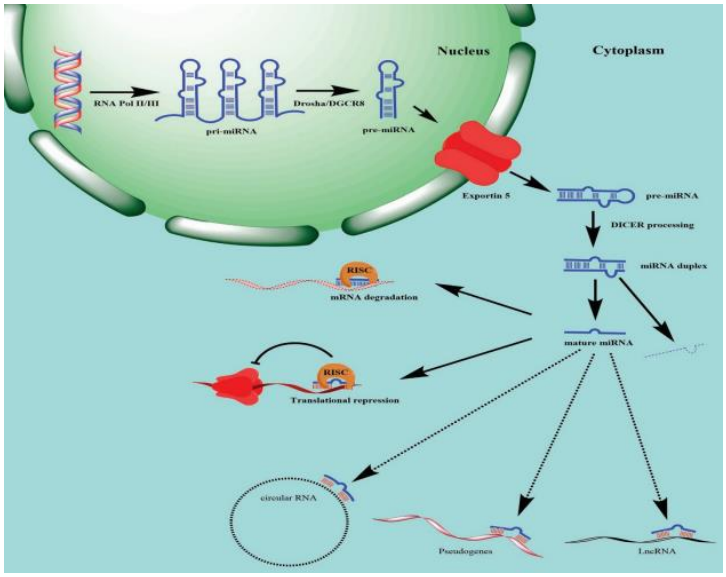

Figure 7 – Schematic of miRNA biogenesis and maturation. The primary miRNA, pri-miRNA is transcribed thanks to RNA polymerase II or III, being then cleaved by Drosha forming pre-miRNA. The latter is exported to the cytoplasm by exportin-5 to be cleaved the loop previously formed in the Dicer, forming a double-stranded structure of miRNA and antisense miRNA (which is degraded). The mature miRNA is then included into miRNA-induced silencing complex (mRISC) which culminates in mRNA degradation or translational repression.
Adapted from Hamam *et al*., 2017 (30)

The dysregulation of miRNAs can be connected to multiple human diseases, one of them being cancer, as consequence of altered miRNA expression because of events like DNA amplification, deletion and mutations related to miRNA *loci*, epigenetic silencing or inhibition of specific miRNA processing. (30) So is important to know firstly the impact that some of the miRNA have in breast cancer situation.

Regarding breast cancer, let-7 miRNA has an abnormal down regulation (normal situations let-7 has overexpression in differentiated epithelial tissues). Like many miRNAs,

let-7 too targets mRNA, more specifically LIN28 mRNA, and causes a regulation via negative feedback, which means that in tumor situations LIN28 proteins have bigger expressions since let-7 is down regulated. Besides this, let-7 is also known to regulate breast cancer tumor initiating cells (T-IC) by targeting HRAS and HMGA. (30)

There is also the miR-200 family which is a family of miRNA that have a tumor suppressor role. The miR-200 family can be divided in two clusters: cluster I composed by three members – miR-200a, miR-200b and miR-429 and cluster II with two members – miR-200c and miR-141. The first cluster can be found on chromosome 1, while the latter, can be found on chromosome 12. MiR-200 family, in terms of breast cancer, is known to intervene in **BMI1 expression regulation** in T-IC by inhibiting zinc-finger E-box binding homeobox ZEB1 and ZEB2 and **suppressing EMT** (epithelial-to-mesenchymal transition), an initiating step in metastasis that is associated with increased breast cancer cell motility and invasiveness. (35–38)

Another important miRNA it's miR-10b which is an oncogenic miRNA (targets HOXD10 and Krüppel-like factor 4 genes) present in metastatic breast cancer being identified up regulated in advanced stages. (39,40)

One other miRNA worth mentioning is miR-21, one of the most over expressed oncogenic miRNAs in breast cancer with an up regulation associated with tumor progression and poor prognosis. This miRNA has the capability of inhibition of tumor-suppressor genes leading to cell growth and invasion, which translates into tumor metastasis. Like the previous mentioned miRNA, so does miR-21 has targets, for example tropomyosin 1α (41–43) and PTEN (promotes MCF-7 breast cancer cell growth). (44)

Contrary to the previous structure, miR-335 tends to silence breast cancer, inhibiting metastasis through targeting the transcription factor Sry-box 4 and extracellular matrix protein tenascin-C. (45,46) This miRNA suppresses the tumors by reducing cell viability and promoting apoptosis. (47)

MicroRNA can circulate free being bound to ribonucleoprotein complexes or high density lipoprotein or they can even be secreted from cells in lipid vesicles, microvesicles, exosomes or apoptic bodies. Through miRNA circulation analysis it is possible to evaluate if there is or not a disease situation, but for that finality it is necessary to detect the circulating miRNAs in the peripheral blood or on other body fluids. (30)

One great characteristic of miRNA relates to their stability and resistance to enzymatic activity from endogenous RNase, giving them the possibility of being used as

diagnostic, prognostic or predictive biomarkers for several diseases, like breast cancer (Table 3). (30)

Table 3 -

Source	No.	MiRNA	Expression level	Diagnostic	Prognostic	Predictive	Validated	Platform
Blood	83	MiR-195, let-7 and -155	Higher in BC patients	Yes	No	No	No	qRT-PCR
Serum	168	MiR-214	Discriminates malignant from benign tumors and healthy subjects	Yes	No	No	No	qRT-PCR
Plasma	247	MiR-127-3p, -376a, -148b, -409-3p, -652 and-801	Higher in BC patients	Yes	No	No	Yes	qRT-PCR
Plasma	137	MiR-148b, -133a, and -409-3p	Higher in BC patients	Yes	No	No	Yes	qRT-PCR
Serum	106	MiR-15a	Higher in BC patients	Yes	No	No	Yes	qRT-PCR
		MiR-18a, -107, -425, -133a, -139-5p, -143, -145, and -365	Lower in BC patients					
Serum	137	MiR-484	Higher in BC patients	Yes	No	No	Yes	qRT-PCR
Serum	1260	MiR-1246, -1307-3p, and 4861-5p	Higher in BC patients	Yes	No	No	Yes	Microarray; qRT-PCR
		MiR-4634 and 4875-5p	Lower in BC patients					
Serum	63	MiR-155, -19a, -181b, and -24	Higher in BC patients	Yes	No	No	No	qRT-PCR
Serum	164	MiR-1, -92a, -133a, and -133b	Higher in BC patients	Yes	No	No	Yes	Microarray; qRT-PCR
Plasma	197	MiR-505-5p, -126b-5p, -21-5p, and -96-5p	Higher in BC patients	Yes	No	No	Yes	qRT-PCR
Serum	90	let-7c	Lower in BC patients	Yes	No	No	No	qRT-PCR
Serum	46	MiR-182	Higher in BC patients	Yes	No	No	No	qRT-PCR
Blood	83	MiR-138	Higher in BC patients	Yes	No	No	No	Microarray; qRT-PCR
Serum	13	MiR-155	Correlates with PR status	Yes	No	No	No	qRT-PCR
Serum	68	MiR-21, -126, -155, -199a, and -335	Associated with histological tumor grade and sex hormone receptor expression	Yes	No	No	No	qRT-PCR
Serum, Plasma	46	MiR-4270, -1225-5p, -1885p, -1202, -4281, -1207-5p, -642b-3p, -1290, and -3141	Higher in BC patients and correlates with stage and molecular subtype	Yes	No	No	Yes	Microarray; qRT-PCR
Serum	102	MiR-202 and let-7b	Higher expression in BC patients and correlates with tumor aggressive and overall survival	Yes	Yes	No	No	qRT-PCR
Serum	87	MiR-148b-3p and 452-3p	Lower in the BC patients	Yes	Yes	No	Yes	ddPCR
		MiR-10b-5p	Higher levels correlate with poor prognosis					
Serum	130	MiR-18b, -103, -107, and -652	Associated with tumor relapse and overall survival in TNBC patients	Yes	Yes	No	Yes	qRT-PCR
Plasma	60	MiR-10b and -373	Higher in breast cancer patients with LN metastasis	Yes	Yes	No	Yes	qRT-PCR
Serum	89	MiR-10b, 34a, and -155	Correlates with tumor stage and/or metastasis	Yes	Yes	No	No	qRT-PCR
Serum	100	miR-29b-2, miR-155, miR -197 and miR -205	Correlates with tumor grade and metastasis	Yes	Yes	No	No	qRT-PCR
Serum	100	MiR-92a	Lower in BC patients, LN metastasis	Yes	Yes	No	No	qRT-PCR
		MiR-21	Higher in BC patients, LN metastasis					
Serum	90	MiR-21-5p, -373, -2055p, and -1944p	Higher in recurrent BC patients	Yes	Yes	No	Yes	qRT-PCR
		MiR-382-5p, -376c-3p, and -411-5p	Lower in recurrent BC patients					
Serum	192	MiR-34a, -93, -373, -17, and -155	Expression correlated with metastasis and HER2, PR, and ER status	Yes	No	No	No	qRT-PCR
Serum	56	miR-125b	Higher expression in non-responsive patients	Yes	No	Yes	No	qRT-PCR
Serum	68	MiR-122	Lower in NR and pCR	No	No	Yes	Yes	DS;
		MiR-375	Higher in NR and pCR					qRT-PCR
Serum	103	MiR-155	Higher in BC patients; decreased level after chemotherapy	Yes	No	Yes	No	qRT-PCR

-Circulating miRNA as biomarkers in diagnostic

Regarding miRNA as biomarkers for diagnostic purposes in breast cancer, in the article "Systemic miRNA-195 differentiates breast cancer from other malignancies and is a potential biomarker for detecting noninvasive and early stage disease", it has been found that some of these structures are quite useful at detecting cancer situations, like let-7a and miR-10b, with the downside of not being completely specific since they are up regulated in multiple kinds of cancer. For a better distinction, it was found that miR-195 could be used to detect more accurately with an high sensitivity and specificity breast cancer, but it was also discovered that if it was evaluated the levels of the three up mentioned miRNA (miR-195, let-7a and miR-155) the sensitivity on breast cancer detection would increase to 94% (comparing to the 88% with only miR-195) (48)

In "Diagnostic potential of PTEN-targeting miR-214 in the blood of breast cancer patients" it was studied potential circulating miRNAs that could target the tumor suppressor PTEN (phosphatase and tensin homologue) through qRT-PCR, and it was observed upper levels of circulating miR-20 and -21 in breast cancer cases, but with the downside of not allowing distinction between benign and malignant tumors. For that it was verified that miR-214 was quite effective. (49)

There have been studied also ways to use miRNA to detect particular stages of breast cancer, specially earlier ones like stage I and II which is very helpful to prevent cases of metastasis. In the particular situation of identifying breast cancer stages I and II there were identified miR-127-3p, miR-148b, miR-409-3p, miR-652 and miR-801 as viable and trustworthy biomarkers, since they have higher levels in these situations. (50) One of the previous miRNA, miR-148b, in conjunction with miR-133a, has been also detected in breast cancer cell lines pointing to a possible tumor origin. (51) Nevertheless, the real origin of the detected circulating miRNAs it is still not yet confirmed and the same applies for the contribution of breast cancer tissue to the miRNA circulating that are identified. For that manner there have been made studies that tried to profile the miRNA expression in breast cancer tumor tissue, with a particular study exposed in "Identification of circulating microRNA signatures for breast cancer detection" identifying miR-1, miR-92a, miR-133a and miR-133b as the most up regulated biomarkers in breast cancer sera (52) or another study, "Tumor microRNA expression profiling identifies circulating microRNAs for early breast cancer detection", that revealed higher levels of miR-505-5p, miR-125b-5p, miR-21-5p and miR-96-5p, in both breast tissue and circulating situations. (53) But unlike most situations

that have been referred previously the change of regulation related to the miRNA, is not always associated with up regulation, and to illustrate that, in “Reduced expression levels of let-7c in human breast cancer patients”, the study demonstrated a down regulation in let-7c in breast cancer tissue, with a later verification resorting to sera of breast cancer patients’ sera. (54)

-Circulating miRNA as biomarkers in prognostic

As it was mentioned initially, circulating miRNA can be applied as biomarkers in prognosis, indicating a possible outcome for a particular patient regarding breast cancer, and not only. For example, in this study “Circulating cell-free cancer-testis MAGE-A RNA, BORIS RNA, let-7b and miR-202 in the blood of patients with breast cancer and benign breast diseases”, both serum levels of let-7b and melanoma-associated antigen-A1, -A2, -A3 and -A12 and CCCTC-binding factor-like mRNA were comparatively higher in invasive breast cancer types than in non-invasive, benign or even healthy situations. Using this values, and comparing to the up regulation of another miRNA, miR-202, it was possible to understand that high levels of this latest micro RNA has a positive correlation with reduced survival. (55)

Another good example is represented in “Diagnostic and prognostic microRNAs in the serum of breast cancer patients measured by droplet digital PCR” where after using a digital PCR to evaluate how miRNAs: miR-10b-5p, miR-145-5p, miR-148b-3p, miR-425-5p and miR-652-3p affected the prognostic in breast cancer it was concluded that an up regulation of miR-10b-5p has an indicative of a poor prognosis in breast cancer patients. (56)

In another article, “A serum microRNA signature predicts tumor relapse and survival in triple-negative breast cancer patients”, it is evaluated the prognostic impact that miRNAs have in triple-negative breast cancer, namely miR-18b, miR-103, miR-107 and miR-652. For that it was used a genome-wide miRNA expression profiling with serum recursion, which lead to the discovery that the four above mentioned miRNA signature was a sign of tumor relapse and overall survival. (57) Equally related to the invasive breast cancer molecular subtype classification, it was realized that higher levels of miR-373 relates to HER2-negative status of the primary tumor, while miR-17 and miR-34a have a correlation with lack of progesteron or estrogen receptors’ status. (58) This miRNA expression profiling has been extremely helpful since it facilitates the establishing of correlations, for example in “The level of circulating miRNA-10b and miRNA-373 in detecting lymph node metastasis of breast cancer: potential biomarkers” it was possible to form a correlation between that profiling and

breast cancer metastasis. In this study it was observed that miR-10b and miR-373 are overexpressed in breast cancer lymph node metastases. (59)

-Circulating miRNA as biomarkers in response treatment prediction

Finally circulating miRNA can also be utilized as predictive biomarkers in breast cancer, the only problem is related to the scarce number of investigations that have been made regarding this subject. For instance, a study used qRT-PCR to investigate miR-155 expression in the sera from individuals with breast cancer, and understood that in breast cancer situation the levels of miR-155 is up regulated, but they also discovered that those same levels would decrease after surgery and chemotherapy (four treatment cycles), which pointed out miR-155 as a possibly good biomarker regarding treatment response/evolution. (60)

Another very interesting article referred that there were made a deep sequencing of circulating miRNAs, using pre-treatment sera from stages II and III locally advanced breast cancer patients who had done neoadjuvant chemotherapy and surgical resection of the tumor. The results obtained allowed to witness that up regulation in miR-122 and down regulation in miR-375, granted the possibility of identify relapsed patients from non-relapsed, results which were latter validated, demonstrating a big correlation high levels of miR-122 and low levels of miR-375, with relapse situations and allowing also a predictive response in regards to chemotherapy treatments. In this same study, it was also identified that elevated levels of miR-375, miR-184, miR-1299 and miR-196a and reduced levels of miR-381, miR-410 and miR-1246 were observed in good responder to neoadjuvant chemotherapy. (61)

All this miRNA structures, and the situations they are used plus how they are detected are synthesized in Table 3.

Limitations in miRNA

Even though the study and analysis of circulating miRNAs as biomarkers in breast cancer diagnosis, prognosis and prediction of treatment response have been showing some promising results, it is a hard process to have reliable and trustworthy biomarkers, since there are some problems associated to miRNA collection and data processing. Probably the biggest problem associated to usage of circulating miRNAs as biomarkers relates to the fact that they are not abundant structures which constitutes a great detection method, as standard profiling techniques such as microarrays become an inadequate option. This will lead to the development of new approaches such as miRNA isolation and enrichment with latter expression profiling. (62,63)

Another problem that has been identified concerns to the way sample selection and processing is made, more specifically choosing between serum and plasma. After studying, it was understood that serum is a more suitable option for sample selection because circulating miRNA are more predominant in serum than in plasma, and also because it prevents the exclusion of large samples as consequence of hemolysis. But serum sampling has a problem, since it can be target of platelet and white blood cell interference during sample preparation. (64)

As it was seen before, qRT-PCR is a very used method for assessing circulating miRNA levels, and even though is very sensitive and less expensive comparatively to other methods, it has a big disadvantage since it can only detect already known circulating miRNA, so important new miRNA that might be great biomarkers but have not been discovered are not detected. (63)

A forth problem identified, relates to appropriate circulating miRNAs housekeeping for normalization of expression levels, which is affected by changes regarding the physiological and pathological status. In order to outline this problem, it has been used equal amounts or either serum or plasma As such, other approaches for normalization have been used, such as using equal amounts of starting material (serum or plasma) (62) or a synthetic spike-in control, which is more reliable than endogenous miRNAs for data normalization. (65)

3. Methodology

3.1. Mass Spectrometry

In regards of mass spectrometry, it is considered a very important technique in terms of structural analysis and uncover of the glycome role. Usually it is used either a matrix-assisted laser desorption/ionization, MALDI (consists in a soft ionization method where the analytes are inserted in organic matrices that are irradiated by a laser; more efficient on proteins than on carbohydrates because the last ones have less ionization efficiency (66)), or an electrospray ionization, ESI (resorts to the use of electrical energy as a way to transfer ions from solution into the gaseous phase; known to be a very sensitive, robust, and reliable method specially adequate to study femto-mole quantities and an alternative form to analyze non-volatile and thermally labile particules; sometimes paired with HPLC (67)). Nonetheless, the most capable instruments for glycan profiling are: modem time-of-flight (TOF), in which

is established a correlation between ion mass-charge and velocity, after applying an initial (equal to all particles) kinetic energy, culminating in an evaluation of the flight time and the measuring the distance between the ions and the detector (68); ion cyclotron resonance (ICR) which allows an analysis with the biggest accuracy resolution and mass measurement wise, comparatively to other types of mass spectrometry, making possible the inclusion of several thousand of particles being tested in a single spectrum (69); and Orbitrap which consists in two electrodes, one central spindle-like (responsible for imprison ions radially) and an exterior barrel-like one, enabling the measurement of mass and charge values from the frequency as result of ionic oscillation without causing any destruction and obtaining as final result a mass spectrum bu using the Furier transforms. (70) These instruments are mostly used because of their high accuracy, working on a scale of subppm, in order to minimize the impact of errors and even, sometimes, additionally it is used to clear false positives, two or more mass spectrometers (Tandem mass spectrometry). In order to increase the accuracy and sensibility of these tests, sample selection and preparation is of the upmost importance, as well as proper selection of glycan fragments (exoglycosidase digestion for example) and exact mass so that the glycome structure can be known. (8)

As it was said in the previous paragraph, one very important step involved is the preparation of glycan fragments, in particular the choice of which method is more adequate to release the target glycan, based on the way it is bonded.

3.1.1. Deglycosylation

For example, N-glycans (as seen before they are attached on asparagine via nitrogen) are released by the enzymatic action of PNGase F (Peptide N-Glicosidade F) and O-glycans (also as seen before they are attached either on serine or threonine via oxygen) are released through reductive beta elimination or ammonia or even sodium borohydride (NaBH4) based beta elimination. As opposed to the first reaction, the last two reactions (cause reduction of each terminal GalNAc residue to its alditol) do not cause a peeling reaction that leads to the cleavage of glycan saccharides and consequently, degraded structures. (71) In case of need of simultaneous release of both N and O-glycan it can be utilized hydrazinolysis, even though the process is not commonly applied since it is a hazardous and difficult procedure. (8)

3.1.2. Purification

The next step is purification, a very important point in the preparation for the analysis capable that will culminate on the separation of glycans and peptides, since the last ones block the signals emitted by glycans, because proteins ionize more easily. Most purification methods are based on the **polarity** difference of these two components, resorting to protein precipitation via methanol, ethanol or acetone (solid phase extraction), for example. Other methods can consist in **lectin affinity enrichment** (glycoprotein concentration, before glycan separation) or use of **amine, amide and graphitized carbon** (for example HPLC) for elution with purpose of glycan retention. Before exposing the samples to analysis they are concentrated either on lyophilizers, via dry nitrogen gas (bubbling) or on vacuum centrifuges. This concentration is fundamental to increase glycan detection (increase sensibility of the method) because of their low level in the solution. One important aspect in this step is glycan preservation, and to do so, solution removal through increased temperatures is not a viable option, unlike the 3 methods previously mentioned. (8)

3.1.3. Glycan Derivatization

Another step that helps improving the results of the mass spectrometry and monitoring of glycan changes between different samples, but on a single test, is the chemical derivatization of glycans (even though it can be made in native state, the choice is a matter of preference), so that the ionization and ion stability increases (decreased fragmentation). This can be achieved by the following reactions: reductive amination, permethylation, metabolic incorporation of stable isotope tags and introduction of isotope labels through enzymatic reactions. (72) It is important to mentioned that some glycans are more prevalent than other (sialylated glycans are the most abundant) and can suppress the signals, observed by mass spectrometry, of lesser abundant glycans (mannose ones for example), problem which can be solved using different ratios of acetonitrile and water in solid phase extraction. (8)

One of the problems adjacent to mass spectrometry glycan analysis it’s that the sources of glycoproteins have salt levels that are high enough to affect negatively the ionization of the analite and in order to solve this problem desalting is necessary, which can be achieved through filtration, dialysis, cation exchange resins or even carbon-based column adsorption. (8)

3.2. Chromatography

In terms of chromatography, as seen before, it is usually paired with mass spectrometry, but nevertheless that are several kinds of chromatographic procedures that can be employed. One example of this analysis method is **Nanoflow Liquid Chromatography** (coupled with capillary electrophoresis and mass spectrometry) which will permit an individual glycan analysis through first separating the target isomers (bigger variety of isomers comparing to the different kinds of composition), so that new glycan markers can be established. In this case, unlike what was referred in the mass spectrometry chapter, derivatization is not mandatory, being possible the N-glycan analysis, for instance, recurring instead to a stationary phase of porous graphitized carbon. (8)

Another technique that sometimes is utilized happens to be **Reversed-phase Chromatography** whose main application is usually glycan separation (in particular isomeric and isobaric structures), and unlike Nanoflow Liquid Chromatography, this procedure requires a derivatization with a hydrophobic tag (higher hydrophobicity leads to a better separation and elution) so that glycans stay on stationary phases (hydrophobic). Another major factor to consider is the contribution of the saccharides in terms of retention which is dependent on their position and linkage. Just like other chromatographic methods this one, should to be paired with either mass spectrometry or UV or Fluorescent detection, even though the first one provides a better resolution in complex mistures with a complete separation of the different glycan species. One of the reasons this kind of chromatography is so advantageous refers to its wide distribution since it is used in almost every bioanalytical research laboratory, facilitating an increase number of facilities investing of glycan analysis. (73)

3.3. Glycan and Lectin Array

Since the discovery of the glycome many tools have been used to push forward the understandings of glycobiology (evolution of glycomics), like mass spectrometry and chromatography mentioned before, but other means have been resorted like arrays, which consists on a interacting partner fixed on a surface in order to detect a certain bindings. In this section the focus will be in **glycan arrays** and **arrays of glycan-binding ligands**, more precisely lectin ones (but it can be also used ligands like antibodies). Glycan and lectin arrays are a revolutionary tool usually coupled with other analytical means and focus on studying

and understanding glycan-mediated recognition involved in signal transduction that occurs in several biological phenomena. (10)

3.3.1. Glycan Array

In regards of **glycan arrays**, these surged as a way to bridge a gap in carbohydrate-binding proteins profiling and are obtained by using a planar surface (interface) to fixate glycans following patterns in an extremely organized way. The enormous advantage of this technique is that, contrary to what happens with classical profiling procedures, it allows a very light usage in terms of material employed, while guaranteeing a greatly thorough analysis with a minimal glycan resource, maximizing the efficiency together with a great sensibility. Glycan arrays have been applied not only to profiling but also in investigation like cancer, viral and bacterial infection and immunity research, demonstrating several applications represented in Figure 6.

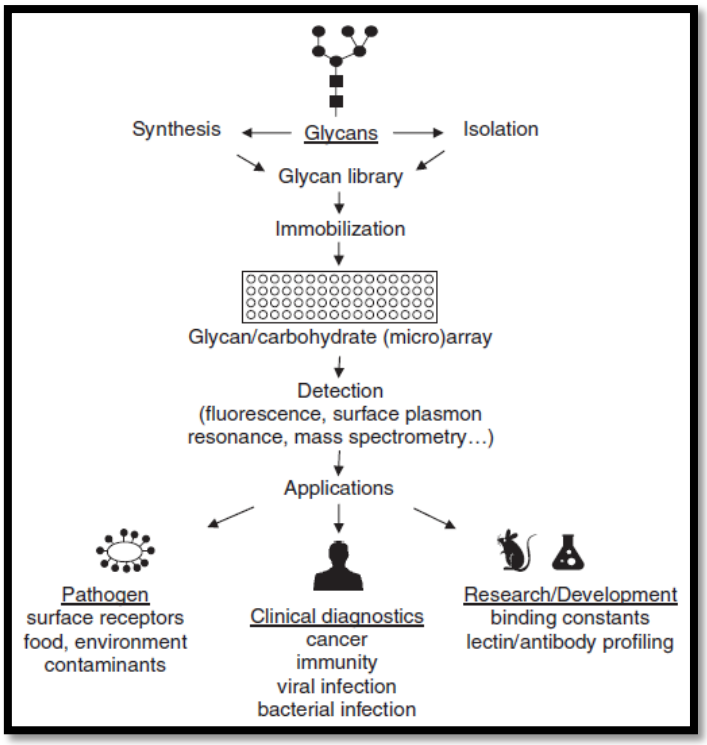

Figure 8 – Schematic of applications of glycan arrays in profiling of glycan-binding ligands Adapted from Katrlík *et al*., 2010 (10)

The glycans introduced in the arrays are synthesized ***in vivo*** through the help of several enzymes like glycosyltransferases, which means the procedure depends heavily on obtaining glycan through tiresome enrichment and **synthesis protocols** or via isolation from

natural sources. This last source of glycan is very important because it contributes to uncover new glycan sequences and their function in biological mechanisms, nevertheless this kind of isolation is quite troublesome in a way that there are a limited number, which confers an obstacle in case of need of large amounts of purified glycans. Regarding the other source, biosynthesis is a process composed by chemical and catalytic steps, which is resorted to when the first source mentioned is not possible, viable or in need of complement. Currently, biosynthesis is still a complex and difficult method to apply due to the complexity associated to the own composition of glycans (great number of saccharides required, connections and branching sites). (10)

The next step is glycan array production. As it was mentioned in the beginning of this segment, glycans have to be fixated on the interface, and thankfully immobilization is a procedure that has been well study well before the interest on glycomics begun, so the real problem will consist on choosing how to fixate the biomolecules. Immobilization requires 3 major points to be followed: completely accessible saccharides chains, oriented immobilization and controlled density of glycans (it affects the affinity to the ligands like antibodies and lectins; the correlaction between glycan density and affinity is not always of direct proportion). Of the several methods the two preferred are: **covalent coupling**, for example, binding to gold surfaces through thiol or more specifically SAMs (self-assembled monolayers of thiol-glycans) that is quite effective because it allows a fixation with control over orientation and surface density (also obtained via hybridization of DNA-glycan in a chip DNA-modified) regarding the recognition element; and **physical adsorption** (for example, adsorption on nitrocellulose/polystyrene-coated glass or immobilization of biotinylated glycans on streptavidin-modified surfaces). The last one, immobilization of biotynilation requires a complete and stable biotinylation of a small percentage of glycans. It's important to mention that in order to conserve their original conformation, glycan usage should be done with their carrier molecules still linked to them. (10)

In terms of detection, the most easy and simple way is optical one, through **fluorescent** and **colorimetric methods**, especially because they are both procedures that have been applied years before in many other experiments. One of the ways to achieve fluorescent labeled glycans is through derivatization using, for instance, **2.6-diaminopyridine as a fluorescent linker**. In case of label-free technology, SPR (Surface Plasmon Resonance) may be a good choice, even though it is more complex, since it allows real time readings with the plus of supplying kinetic and affinity constants of interactions (also used in glycan

recognition by plant lectin). This technique can be also improved by coupling mass spectrometry. (10)

3.3.2. Lectin Array

Reversibly to what happens in glycan arrays, lectin arrays or microarrays are used in profiling glycans (capable of detecting the quality and quantity of several epitopes) or even in helping, uncover glycosylation patterns specific of certain diseases (breast cancer for example), with the downsides that are certain glycans that cannot be recognized because of lack of correspondence between these two and also of promiscuous binding regarding some lectins. Besides this disadvantage, lectin arrays present a characteristic very unique non existing in other classic methods, which is an analysis on intact structures, even in complete cells. (10)

The lectin microarrays analysis method grants the possibility of global glycomic profiling of two kinds of glycans: N- and O- glycan, basing the detection on lectin and carbohydrate interactions with an high sensibility, without being complex or very time consuming. This technology surged as complement of previous methods employed such as mass spectrometry, known for its more complex steps since it requires more intricate steps regarding sample preparation and experiment conditions optimization. (74)

There are three principle types of lectin arrays that are depicted in figure 7: the direct assay where the lectins are fixated on a chip's surface, followed by the injection of the labeled sample (glycoproteins or whole cell) over that same surface; the glycoprotein-lectin array where the opposite of the previous array happens, which translates to fixation of prepurified samples on the chip's surface followed by injection of labeled lectins; the sandwich assay where besides lectins are also used antibodies, being the latest immobilized on the chip's surface, with an injection of a sample recognized by those, as the next step, which is followed with a final injection of labeled lectins. (10)

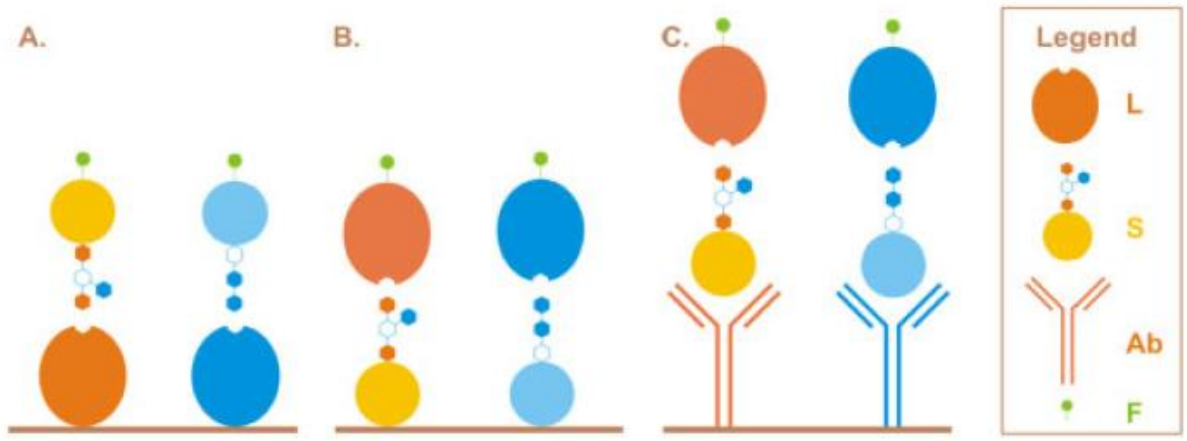

Figure 9 - Depiction of the three principle types of lectin arrays. Direct detection (A) involves lectin immobilization on the array's surface with a latter injection of labeled sample over the surface. In the Reverse Phase Sample Blot (B) the sample is immobilized in the surface while labeled lectin is posteriorly injected. Sandwich Detection (C) consists in an immobilized antibody that will recognize the sample, with a latter injection of a labeled lectin. Abbreviations in the figure: L – Lectin; S- Sample (glycoproteins or cells for example); Ab – Antibody; F – Fluorescent dye
Adapted from Katrlík *et al.*, 2010 (10)

Most of the lectins used in arrays or microarrays are plant based and present an extremely importance regarding biomarker research (explained in more detail in the following segments), but more sources can be used, for example human, animal, fungus or microbial ones. (10) The fact that lectins are naturally obtained (via chromatographic so that they can be isolated) constitutes a problem that can lead to variations between batches which can be solved through recombinant DNA technology to select the preferred characteristics in those proteins so that they can recognized the target glycans. (75)

Regarding the fixation procedure in lectin arrays, it is done in a similar way like it was mentioned before, nevertheless the most chosen method of fixation requires lectin immobilization on glass slides that can be modified in several ways such as epoxy or aldehyde groups. As an immobilization platform, PDMS (polymethylsiloxane) can be used for covalent grafting of lectins and if possible, is preferred that this step be done, not randomly, but in the most oriented way in order to improve stability and binding activity of these proteins. These factors plus density control have a great impact on the effectiveness of the arrays which can be reached with nanotechnology or peptide borono lectins. (10)

As a starting point, lectin arrays should have a bigger number of lectins on the array's surface (normally up to 94), to discover the optimal combination of lectins for the target glycan. After this is achieved in the following assays the number of lectins is decreased (between 3 and 20). (10)

The detection procedures for lectin arrays, like glycan ones, can be of two types: labeled (fluorescent based in most cases) or label free. As it was mentioned before, fluorescent detection is more commonly used because it is a very well-known and well-studied method besides also being a quite simple and robust way of gathering data, even though label free methods have the advantage of less probability of changing the bio recognition pattern of the target glycoprotein. (10)

So as it was demonstrate these methods are extremely important in glycan (and not only) profiling, but too are very useful for biomarker discovery in terms of disease screening.

4. Lectins

Lectins are a large group of proteins which are well known for their capability of erythrocytes agglutination (reason why they are also called agglutinins), but are also considered a "decoding system" since they are receptors specific for glycans, which can be found widely dispersed in nature, especially in plants (plant lectins constitute the largest and most studied type of these proteins) and in their vegetative organs such as roots, leaves, flowers, fruits, nectar and mostly in seeds. (7)

4.1. Carbohydrate Recognition Domain

As plant lectins are the most studied within the existing collective, the focus for now will be on them. These type of lectins, while glycan receptors, are non-immune proteins capable of recognizing and binding carbohydrates structures, in either a monospecific or polyspecific way. (76) The binding happens thanks to the existence of a **Carbohydrate Recognition Domain (CRD)** which exists within the lectins structure. (77) Even though currently, plant lectins are the ones being explained in this segment, the existence of CRD not only is not limited to plant lectins as it is present in all lectins, constituting a small section of their structure, named "loop", with no more than 200 amino acids. CRD structures are also known to have cavities on their surface, which can be shallow or deeper impacting negatively or not the way the binding is done with saccharides regarding affinity, being that usually the preferred type are the deeper ones since they grant an higher affinity binding. Besides the existence of deeper cavities to improve the binding affinity, it is also ideal to have multiple association of CRD in the same lectin. The length of the polypeptide chain of the CRD is important at the level of interaction number in the domain. Also the water molecules and

bivalent cations Ca2+ and Mn2+ that were binding to lateral chains of amino acid residues of lectin are crucial. Another characteristic of CRD rests in the fact that CRDs with no structural relation can recognize the same saccharides, demonstrating a certain level or redundancy, but on the other side, structurally similar CRDs may recognize different oligosaccharides. (78–80)

Consequently the demonstration of these characteristics has resulted in plant lectins being used as tools in cancer research, for example in differentiation between benign and malignant tumors or in the evaluation of glycosylation in metastasis cases by decoding the cellular glycome. (76)

Plant lectins are consumed constantly by living beings on a daily basis and are known to have both biological and chemical analog properties, one of those being the capacity to resist to digestion. This characteristic proves to be anti-nutritive or even, toxic in high quantities, because lectins bind to glycan receptors in the digestive tract, which can lead to all types of injuries, starting from lesions in epithelium membranes or disruption of nutrient digestion and absorption to major ones like, modified metabolism of macronutrients or damaged tissues or organs. (7)

4.2. Classification

One of the problems associated to plant lectins is related to their diversity that leads to a difficulty in establishing a classification, so as a consequence there several ones based on miscellaneous factors, of which are worth mentioning the following: (81)

4.2.1. Carbohydrate-binding classification

- Mannose- and mannose/glucose-
- Mannose/maltose-
- Gal/GalNAc-
- N-acetylglucosamine (GlcNAc)/(GlcNAc)-
- Fucose-
- Sialic acid-biding lectins

4.2.2. Domain architecture-based classification

- Merolectins (possess only one CRD, which results on the lack of ability of cell agglutination)
- Hololectins (possess at least two homologous CRD, having the ability of cell agglutination or glycoconjugates precipitation)
- Superlectins (possess at least two non-homologous CRD, capable of recognizing carbohydrates which have no structural relation)
- Chimerolectins (possess at least one CRD and characterized for having independent biological activity)

Regarding this last classification, of the several groups on which plant lectins were divided (Figure 9), the one most isolated and studied are hololectins. (7)

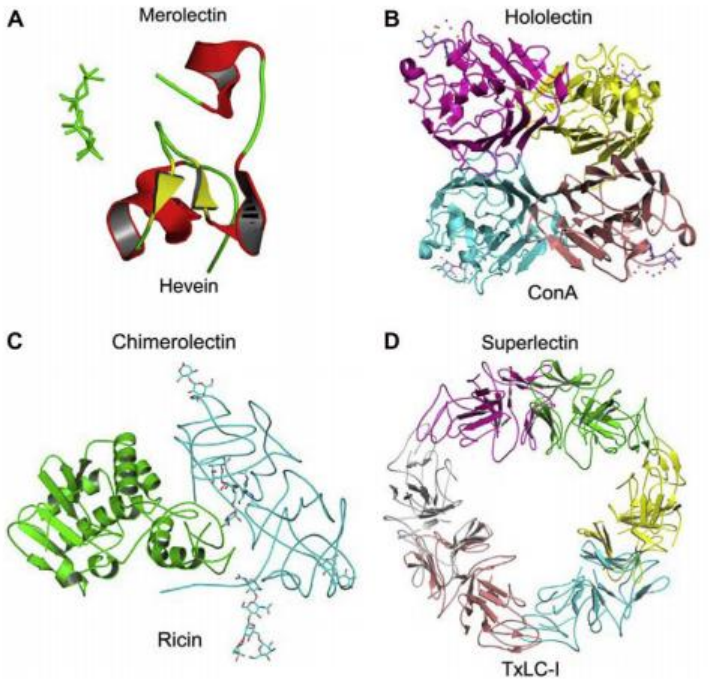

Figure 10 - Schematic representation of merolectins, hololectins, chimerolectins, and superlectins. The four types represented here are based on the overall structure of mature plant lectins. Merolectins (A) consist on a single carbohydrate-binding domain – Hevein; Hololectines (B) only composed by carbohydrate-bindgin domains, with a minimal of two identical or homologous domains that bind to the same or identical sugars – ConA; Chimerolectins (C) result from the fusion between proteins with one or more carbohydrate-binding domains tandemly arrayed to a non-related domain – Ricin; Superlectins (D) have only two carbohydrate-binding domains – TXCL-1.
Adapted from Liu *et al*., 2010 (76)

4.2.3. Van Damme classification – 2008

Table 4 – Actualized version of plant lectin based classification according to Van Damme *et al.*,2008. This classification takes into consideration sequence similarities and serological relatedness and also sequence and structural homology.

	Lectin Family	Typical Saccharide Ligands	Predicted Localization
Plant	*Agaricus bisporus* lectin family	GlcNAc/GalNAc, galactose	Nucleus, cytosol
	Amaranthin family	GalNAc	Nucleus, cytosol
	Chitinase-related agglutinin family	High-mannose *N*-glycans	Vacuole, membrane bound
	Cyanovirin family	Mannose	Nucleus
	Euonymus europaeus lectin family	Galactosides, high-mannose *N*-glycans	Nucleus, cytosol
	Galanthus nivalis lectin family	Mannose	Vacuole, nucleus, cytosol, or membrane bound
	Hevein family	Chitin	Vacuole
	Jacalin family	Mannose- and galactose-specific subgroup	Nucleus, cytosol, vacuole
	Legume family	Mannose	Vacuole, nucleus, cytosol, or membrane bound
	LysM family	Chitin, peptidoglycan	Vacuole, nucleus, cytosol, or membrane bound
	Nicotiana tabacum lectin family	(GlcNAc)n, high-mannose and complex *N*-glycans	Nucleus, cytosol
	Ricin-B family	Gal/GalNAc, sialylated Gal/GalNAc	Vacuole, nucleus, cytosol

Adapted from Van Damme *et. Al, 2008* (82)

4.3. Role of Lectins on the detection of Biomarkers

Lectins, as mentioned above, are a very important and increasingly used tool for detection of biomarkers, regarding glycan profiling and consequently aberrant glycosylation identification and description. Usually lectins are not used alone, but in conjunction with

other methods like, HPLC or mass spectrometry which allow the study and detection of specific alterations in terms of glycans, through the separation of oligosaccharides released from glycoproteins, as was studied at the first chapters of this article. Monoclonal antibodies and lectins can be also employed together for a more accurate detection of aberrant glycosylation, allowing the quantification of the proportion regarding the detected alterations. (21)

In order to demonstrate real examples of employed lectins on biomarker detection, it will be used as reference, the ones applied in the cases mentioned at the biomarkers chapter.

4.3.1. CA15-3 (MUC1)

In this case, the method chosen was taken from "Use of CA15-3 for screening breast cancer: An antibody-lectin sandwich assay for detecting glycosylation of CA15-3 in sera" and it is an antibody-lectin sandwich assay to detect CA15-3 (MUC1) glycosylation, where the mouse antibody anti-CA15-3 monoclonal antibody is immobilized, following up the capture of CA15-3 present in serum. In the article taken as reference it is used also detection via ELISA to measure CA15-3 levels in serum but it was concluded that while this first method allowed serum dilutions of 1:50–1:100, the antibody-lectin sandwich assay detects serum dilutions that exceeds 1:2,000, meaning that the antibody-lectin sandwich assay is much more sensitive than the ELISA, probably because of the increased signal as consequence of the abundant glycosylation of CA15-3, and also allows achieving linear results. (83)

The antibody-lectin sandwich assay for the detection of CA15-3 (MUC1) glycosylation started with a coating of immunoplates (96 wells) with mouse anti-CA15-3 monoclonal antibody in PBS (phosphate-buffered saline) for 16 hours at 4°C. This step was then followed by the addition of FBS (fetal bovine serum) in TBS-T (Tris-buffered saline containing 1% Tween-20) for 3h at room temperature. The FBS, which contains high concentration of glycoprotein, will act as blocking agent which will block non-specific reactions of human serum protein. After this, the FBS and the monoclonal antibody coating is oxidized with sodium metaperiodate, with the finality of block lectin binding, followed then by a second oxidization with 1% oxidized bovine serum albumin (oBSA). The next step was to dissolve BSA and remove the sodium metaperiodate, to then apply the diluted sera (1:3000 or 1:4000) in the wells for 2h at 37°C. Biotinylated lectins were prepared after at a concentration of 1 µg/ml in TBS-T in order to by applied and react in the wells for 1 h at room temperature, which would lead to lectin bound to the CA15-3 glycan. The detection of

CA15-3 (MUC1) was done with poly-HRP-conjugated streptavidin which would culminate in color detection after TBS-T washes (Figure 11). (83)

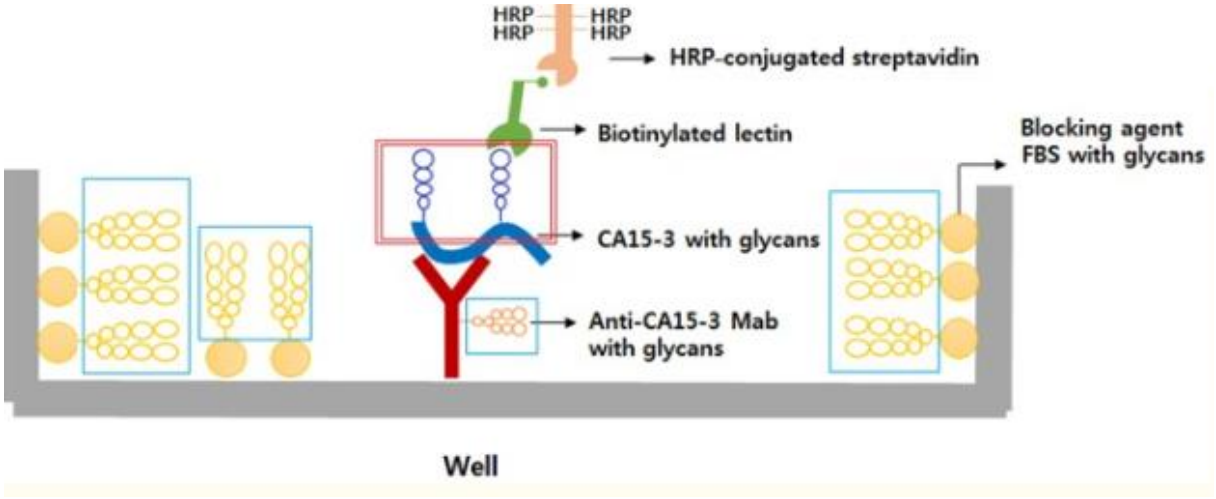

Figure 11 – Schematic diagram of the antibody-lectin sandwich assay for detecting glycosylation of serum CA15-3. After immobilizing anti-CA15-3 antibody, this one will capture CA15-3 in serum, allowing the detection of CA15-3 with a biotinylated lectin followed by HRP-conjugated streptavidin. Blues boxes represent glycosyls of the coated anti-CA15-3 monoclonal antibody, and of FBS used as a blocking agent. The red boxes represent glycans of CA15-3.
Adapted from Choi *et al*., 2018 (83)

This assay was done using twelve different types of lectins in order to study the lectin-binding properties of the coating antibody and FBS (Table 5). Of those twelve only ConA exhibited strong reactivity towards CA15-3, which may be due to the other lectins having more difficulty on the access to this mucin, probably because of the hindering caused by the glycoprotein in question forming aggregates or gels or even because of the existence of a self-aggregation domain in CA15-3. As was mentioned before CA15-3 is also extremely O-glycosylated which also increases the difficulty of accessing it with lectins (increased access can be made with perchloric acid treatment). (83)

Table 5 – Binding properties of lectins used in "Use of CA15-3 for screening breast cancer: An antibody-lection sandwich assay for detectin glycosylation of Ca15-3 in sera"

Abbreviation	Full name	Specificity	Type of glycan linkage
ConA	Concanavalin A	α-linked mannose high mannose type glycans	N-linked
SNA	*Sambucus nigra* lectin	Sialic acid attached to galactose in α-2,6 linkage	N- and O-linked
PNA	Peanut agglutinin	Galactose attached to *N*-acetylgalactosamine in β-1,3 linkage	O-linked
RCA I	*Ricinus communis* agglutinin I	Galactose attached to *N*-acetylglucosamine in β-1,4 linkage	N-linked
HPA	*Helix pomatia* agglutinin	*N*-acetylgalactosamine	O-linked
UEA-I	*Ulex europaeus* agglutinin I	α-linked fucose	N- and O-linked
AAL	*Aleuria aurantia* lectin	Fucose linked (α −1,6) to *N*-acetylglucosamine or fucose linked (α −1,3) to *N*-acetyllactosamine	N- and O-linked
LCA	*Lens culinaris* agglutinin	α-linked mannose α-linked fucose attached to the *N*-acetylchitobiose portion markedly enhances affinity	N-linked
AOL	*Aspergillus oryzae* lectin	α1,6-linked fucose	N-linked
WFA	*Wisteria floribunda* Lectin	Carbohydrate structures terminating in *N*-acetylgalactosamine linked α or β to the 3 or 6 position of galactose	NA
MAA	*Maackia amurensis* lectin II	Sialic acid attached to galactose in α-2,3 linkage	N- and O-linked
Jacalin	Jackfruit (*Artocarpus heterophyllus*)	Galactose attached to *N*-acetylgalactosamine in β-1,3 linkage/mono- or disialylated form of *O*-glycan	O-linked

NA, not available.

Adapted from Choi *et al*., 2018 (83)

In order to determine the values of the anti-CA15-3-ConA sandwich assays it was made a subtraction operation between optical density of the full series of reactions and the optical density of the mirror plate reaction (which, in comparison to the assay above described, has the coating antibody reaction omitted in order to remove non-specific reactions of the serum from the equation).

4.3.2. IgA1

In this case, it was used as reference the article "Serum IgA1 shows increased levels of a2,6-linked sialic acid in breast cancer", where it was studied how IgA1 could be used in breast cancer screening and as a biomarker. To understand that, it was measured the levels of IgA1 in the serum of breast cancer patients in order to try to uncover an aberrant glycosylation in IgA1 by comparison with serum from healthy women. Through ELLA (Enzyme-linked Lectin Assay) and ELISA assays, and also with mass spectrometry it was possible to comprehend how this type of cancer affected IgA1. Firstly, it was used the ELISA assay in order to access antibody IgA1 glycosylation, but the ELISA method was changed (to ELLA assay) to include *Helix pomatia* agglutinin (HPA), *sambucus nigra* (SNA) and *Maackia amurensis* lectin II (MAL-II) as detection reagents instead of an antibody. The first step includes coating plates with IgA1 diluted in PBS for 1h at 378C. This is then followed by 3 washes with PBST, making the plates blocked. After that it is applied 50 mL of serum sample (diluted 1:5, 1: 10 or 1:20 for the HPA ELLA and 1: 32 000 in for the SNA/MAL-II ELLA) in the wells for 30 min at 378C, with another wash in the end accompanied with the

addition of HPA-biotin for 1h at room temperature, followed by another washing with PBST and appliance of poly-streptavidin-HRP to each well for 1 h. Finally the plate was washed again with PBST, then three times with distilled water and for conclusion it was added TMB substrate each well for 1– 15 min. The reaction in the end is quenched allowing to the absorbance to be measured. Initially, it was used a pooled serum as reference in the initial HPA and the SNA/MAL-II ELLA, which was later substituted by a a standard using IgA from human colostrum for the later HPA ELLA. According to the initial ELISA and ELLA analysis, HPA bound to the antibody IgA1 in a more extensive way in breast cancer situation than to IgA from individuals devoid of this pathology. The lectin HPA has a nominal binding specificity for GalNAc (N-acetylgalactosamine), and even though it also recognizes Forssman antigen, blood group A antigen, Tn antigen, sialic acid (SA) and GlcNAc (N-acetylyglucosamine), the nominal binding specificity of HPA is very important, since in this study was determined that the IgA1 in the serum of breast cancer patients is GalNAcylated and for that reason it is possible to understand why there is a more extensive bound between HPA and IgA1 in breast cancer situations. (29)

The incorporation of the lectins SNA and MAL-II in ELLA assay had as purpose the evaluation of there was or not increased values of α 2,6- and α 2,3-linked sialic acid in breast cancer (sialylation occurence). During these tests it was observed an increased binding level in SNA lectin in the IgA1 captured from breast cancer patients. Since via western blot analysis was verified that SNA bound to both the heavy and light chains of IgA1 from breast cancer preparations, the increased sialylation values come probably from the α 2,6-linked sialic acid. (29)

And secondly, using mass spectrometry (with a previous IgA1 purification followed by a release of its glycans) and structural analysis it was observable an increase in the TF antigen in the O-glycans and sialylation of the N-glycans (even though IgA1 from both breast cancer patients and normal individuals exhibited similar N-glycans) in the serum IgA1 from patients with breast cancer.(29)

Through these results is possible to understand that IgA1 is a very useful biomarker for a prognosis situation.

This study shows that the IgA1 in the serum of BCa patients is GalNAcylated and that it also has a different, sialylated, repertoire.

5. Discussion and Conclusions

The analysis and study of glycome has evolved a great deal over the last few decades, however it has not been easy to study and much has yet to be discovered due to the complexity underlying the glycosylation process and the fact that glycan structure is formed and rapidly restructured without any genomic alteration.

The association between lectins and glycans as a means of detecting breast tumors has been increasingly investigated in recent years, as it allows the identification of glycomic aberrations by comparing changes between glycan structures to identify differences (affected glycosylation). However, as it has been seen earlier, altered glycosylation does not always imply the occurrence of a cancer-associated pathological condition, since by augmentation of a biomarker it does not always refers to the existence of a tumor, as seen in the above mentioned case of IgA1 where its level may be increased because of circulating B lymphocytes or where it is not possible to (sometimes) clearly distinguish the evolutionary state of the disease. This problem then leads to a desire to increase the specificity and sensitivity of the analyzes, which according to the research carried out, concludes that all investigators agree that the glycan and lectin screening method should be complemented with others such as mass spectrometry or chromatography, or additional use of antibodies.

Another problem that has been found is that while many biomarkers are being exploited as a way to be used for detection, few have actually been authorized by authorities such as the FDA because many of them are not as effective as the authorized ones due to the difficulty and complexity that is associated with the study and profiling of human glycome for the reasons already mentioned.

This monography gives a brief glimpse into a huge universe of glycome research, including examples of how it can be done and what advantages it can bring to human health by knowing its role in the functioning of the human organism at the cellular level, focusing breast cancer, but although it has very promising results, it is an area that still needs to be thoroughly studied until it can be applied on a larger scale, as it still has some unresolved issues to be completely reliable, without forgetting that it has already opened many doors regarding the way cells interact with each other and with other molecules.

6. Bibliography

1. Tao ZQ, Shi A, Lu C, Song T, Zhang Z, Zhao J. Breast Cancer: Epidemiology and Etiology. Cell Biochem Biophys. 2015;
2. Sun YS, Zhao Z, Yang ZN, Xu F, Lu HJ, Zhu ZY, et al. Risk factors and preventions of breast cancer. International Journal of Biological Sciences. 2017.
3. Ghoncheh M, Pournamdar Z, Salehiniya H. Incidence and mortality and epidemiology of breast cancer in the world. Asian Pacific J Cancer Prev. 2016;
4. Society AC. Breast Cancer Facts and Figures 2017-2018. Atlanta: American Cancer Society, Inc. 2017.
5. Cheang MCU, Martin M, Nielsen TO, Prat A, Voduc D, Rodriguez-Lescure A, et al. Defining Breast Cancer Intrinsic Subtypes by Quantitative Receptor Expression. Oncologist. 2015;
6. World Cancer Research Fund/American Institute for Cancer Research. Diet, nutrition, physical activity and breast cancer. Contin Updat Proj Expert Rep 2018. 2018;
7. Ribeiro AC, Ferreira R, Freitas R. Plant Lectins: Bioactivities and Bioapplications. In: Studies in Natural Products Chemistry. 2018.
8. An HJ, Kronewitter SR, de Leoz MLA, Lebrilla CB. Glycomics and disease markers. Current Opinion in Chemical Biology. 2009.
9. Taron CH, Rudd PM. Glycomics : A rapidly evolving field with a sweet future Technical Advances in Glycomics [Internet]. 1982 [cited 2019 Sep 13]. Available from: https://www.neb.com/tools-and-resources/feature-articles/glycomics-a-rapidly-evolving-field-with-a-sweet-future
10. Katrlík J, Švitel J, Gemeiner P, Kožár T, Tkac J. Glycan and lectin microarrays for glycomics and medicinal applications. Medicinal Research Reviews. 2010.
11. Corfield A. Eukaryotic protein glycosylation: a primer for histochemists and cell biologists. Histochemistry and Cell Biology. 2017.
12. Wi GR, Moon BI, Kim HJ, Lim W, Lee A, Lee JW, et al. A lectin-based approach to detecting carcinogenesis in breast tissue. Oncol Lett. 2016;
13. Varki A, Cummings RD, Esko JD, Stanley P, Hart GW, Aebi M, et al. Essentials of glycobiology, third edition. Cold Spring Harbor Laboratory Press. 2017.
14. Wang P, Wang H, Gai J, Tian X, Zhang X, Lv Y, et al. Evolution of protein N-glycosylation process in Golgi apparatus which shapes diversity of protein N-glycan

structures in plants, animals and fungi. Sci Rep. 2017;

15. Lodish H, Berk A, Zipursky S, Al E. Glycosylation in the ER and Golgi Complex. Mol Cell Biol. 2000;
16. Pinto R, Carvalho AS, Conze T, Magalhães A, Picco G, Burchell JM, et al. Identification of new cancer biomarkers based on aberrant mucin glycoforms by in situ proximity ligation. J Cell Mol Med. 2012;
17. Tran DT, Ten Hagen KG. Mucin-type o-glycosylation during development. Journal of Biological Chemistry. 2013.
18. Hanisch FG. O-glycosylation of the mucin type. Biological Chemistry. 2001.
19. Ten Hagen KG, Fritz TA, Tabak LA. All in the family: The UDP-GalNAc:polypeptide N-acetylgalactosaminyltransferases. Vol. 13, Glycobiology. 2003.
20. Li Q, Li G, Zhou Y, Zhang X, Sun M, Jiang H, et al. Comprehensive N-Glycome Profiling of Cells and Tissues for Breast Cancer Diagnosis. J Proteome Res [Internet]. 2019 Jun 7 [cited 2019 Oct 19];18(6):2559–70. Available from: http://www.ncbi.nlm.nih.gov/pubmed/30889355
21. Kirwan A, Utratna M, O'Dwyer ME, Joshi L, Kilcoyne M. Glycosylation-Based Serum Biomarkers for Cancer Diagnostics and Prognostics. BioMed Research International. 2015.
22. Hakomori S. Glycosylation defining cancer malignancy: New wine in an old bottle. Proceedings of the National Academy of Sciences of the United States of America. 2002.
23. Strimbu K, Tavel JA. What are biomarkers? Vol. 5, Current Opinion in HIV and AIDS. 2010. p. 463–6.
24. Nath S, Mukherjee P. MUC1: A multifaceted oncoprotein with a key role in cancer progression. Trends in Molecular Medicine. 2014.
25. Carson DD. The cytoplasmic tail of MUC1: A very busy place. Science Signaling. 2008.
26. Terävä J, Tiainen L, Lamminmäki U, Kellokumpu-Lehtinen P-L, Pettersson K, Gidwani K. Lectin nanoparticle assays for detecting breast cancer-associated glycovariants of cancer antigen 15-3 (CA15-3) in human plasma. PLoS One. 2019;
27. Gomes MM, Wall SB, Takahashi K, Novak J, Renfrow MB, Herr AB. Analysis of IgA1 N-glycosylation and its contribution to FcαRI binding. Biochemistry. 2008;
28. Stuchlová Horynová M, Raška M, Clausen H, Novak J. Aberrant O-glycosylation and anti-glycan antibodies in an autoimmune disease IgA nephropathy and breast

adenocarcinoma. Cellular and Molecular Life Sciences. 2013.

29. Lomax-Browne HJ, Robertson C, Antonopoulos A, Leathem AJC, Haslam SM, Dell A, et al. Serum IgA1 shows increased levels of a2,6-linked sialic acid in breast cancer. Interface Focus. 2019;
30. Hamam R, Hamam D, Alsaleh KA, Kassem M, Zaher W, Alfayez M, et al. Circulating microRNAs in breast cancer: novel diagnostic and prognostic biomarkers. Cell death & disease. 2017.
31. Hammond SM. An overview of microRNAs. Advanced Drug Delivery Reviews. 2015.
32. Davis BN, Hata A. Regulation of MicroRNA Biogenesis: A miRiad of mechanisms. Cell Communication and Signaling. 2009.
33. Carthew RW, Sontheimer EJ. Origins and Mechanisms of miRNAs and siRNAs. Cell. 2009.
34. Pradhan AK, Emdad L, Das SK, Sarkar D, Fisher PB. The Enigma of miRNA Regulation in Cancer. In: Advances in Cancer Research. 2017.
35. Huber MA, Azoitei N, Baumann B, Grünert S, Sommer A, Pehamberger H, et al. NF-κB is essential for epithelial-mesenchymal transition and metastasis in a model of breast cancer progression. J Clin Invest. 2004;
36. Yang J, Mani SA, Donaher JL, Ramaswamy S, Itzykson RA, Come C, et al. Twist, a master regulator of morphogenesis, plays an essential role in tumor metastasis. Cell. 2004;
37. Shimono Y, Zabala M, Cho RW, Lobo N, Dalerba P, Qian D, et al. Downregulation of miRNA-200c Links Breast Cancer Stem Cells with Normal Stem Cells. Cell. 2009;
38. Korpal M, Lee ES, Hu G, Kang Y. The miR-200 family inhibits epithelial-mesenchymal transition and cancer cell migration by direct targeting of E-cadherin transcriptional repressors ZEB1 and ZEB2. J Biol Chem. 2008;
39. Ma L, Teruya-Feldstein J, Weinberg RA. Tumour invasion and metastasis initiated by microRNA-10b in breast cancer. Nature. 2007;
40. Rowland BD, Peeper DS. KLF4, p21 and context-dependent opposing forces in cancer. Nature Reviews Cancer. 2006.
41. Frankel LB, Christoffersen NR, Jacobsen A, Lindow M, Krogh A, Lund AH. Programmed cell death 4 (PDCD4) is an important functional target of the microRNA miR-21 in breast cancer cells. J Biol Chem. 2008;
42. Zhu S, Si ML, Wu H, Mo YY. MicroRNA-21 targets the tumor suppressor gene tropomyosin 1 (TPM1). J Biol Chem. 2007;

43. Qian B, Katsaros D, Lu L, Preti M, Durando A, Arisio R, et al. High miR-21 expression in breast cancer associated with poor disease-free survival in early stage disease and high TGF-β1. Breast Cancer Res Treat. 2009;
44. Si ML, Zhu S, Wu H, Lu Z, Wu F, Mo YY. miR-21-mediated tumor growth. Oncogene. 2007;
45. Png KJ, Yoshida M, Zhang XHF, Shu W, Lee H, Rimner A, et al. MicroRNA-335 inhibits tumor reinitiation and is silenced through genetic and epigenetic mechanisms in human breast cancer. Genes Dev. 2011;
46. Tavazoie SF, Alarcón C, Oskarsson T, Padua D, Wang Q, Bos PD, et al. Endogenous human microRNAs that suppress breast cancer metastasis. Nature. 2008;
47. Heyn H, Engelmann M, Schreek S, Ahrens P, Lehmann U, Kreipe H, et al. MicroRNA miR-335 is crucial for the BRCA1 regulatory cascade in breast cancer development. Int J Cancer. 2011;
48. Heneghan HM, Miller N, Kelly R, Newell J, Kerin MJ. Systemic miRNA-195 Differentiates Breast Cancer from Other Malignancies and Is a Potential Biomarker for Detecting Noninvasive and Early Stage Disease. Oncologist. 2010;
49. Schwarzenbach H, Milde-Langosch K, Steinbach B, Müller V, Pantel K. Diagnostic potential of PTEN-targeting miR-214 in the blood of breast cancer patients. Breast Cancer Res Treat. 2012;
50. Cuk K, Zucknick M, Madhavan D, Schott S, Golatta M, Heil J, et al. Plasma MicroRNA Panel for Minimally Invasive Detection of Breast Cancer. PLoS One. 2013;
51. Shen J, Hu Q, Schrauder M, Yan L, Wang D, Medico L, et al. Circulating miR-148b and miR-133a as biomarkers for breast cancer detection. Oncotarget. 2014;
52. Chan M, Liaw CS, Ji SM, Tan HH, Wong CY, Thike AA, et al. Identification of circulating microRNA signatures for breast cancer detection. Clin Cancer Res. 2013;
53. Matamala N, Vargas MT, González-Cámpora R, Miñambres R, Arias J, Menéndez P, et al. Tumor MicroRNA expression profiling identifies circulating MicroRNAs for early breast cancer detection. Clin Chem. 2015;
54. Li XX, Gao SY, Wang PY, Zhou X, Li YJ, Yu Y, et al. Reduced expression levels of let-7c in human breast cancer patients. Oncol Lett. 2015;
55. Joosse SA, Müller V, Steinbach B, Pantel K, Schwarzenbach H. Circulating cell-free cancer-testis MAGE-A RNA, BORIS RNA, let-7b and miR-202 in the blood of patients with breast cancer and benign breast diseases. Br J Cancer. 2014;
56. Mangolini A, Ferracin M, Zanzi MV, Saccenti E, Ebnaof SO, Poma VV, et al.

Diagnostic and prognostic microRNAs in the serum of breast cancer patients measured by droplet digital PCR. Biomark Res. 2015;

57. Sahlberg KK, Bottai G, Naume B, Burwinkel B, Calin GA, Børresen-Dale AL, et al. A serum MicroRNA signature predicts tumor relapse and survival in triple-negative breast cancer patients. Clin Cancer Res. 2015;
58. Eichelser C, Flesch-Janys D, Chang-Claude J, Pantel K, Schwarzenbach H. Deregulated serum concentrations of circulating cell-free microRNAs miR-17, miR-34a, miR-155, and miR-373 in human breast cancer development and progression. Clin Chem. 2013;
59. Chen W, Cai F, Zhang B, Barekati Z, Zhong XY. The level of circulating miRNA-10b and miRNA-373 in detecting lymph node metastasis of breast cancer: Potential biomarkers. Tumor Biol. 2013;
60. Sun Y, Wang M, Lin G, Sun S, Li X, Qi J, et al. Serum MicroRNA-155 as a Potential Biomarker to Track Disease in Breast Cancer. PLoS One. 2012;
61. Wu X, Somlo G, Yu Y, Palomares MR, Li AX, Zhou W, et al. De novo sequencing of circulating miRNAs identifies novel markers predicting clinical outcome of locally advanced breast cancer. J Transl Med. 2012;
62. Hamam R, Ali AM, Alsaleh KA, Kassem M, Alfayez M, Aldahmash A, et al. microRNA expression profiling on individual breast cancer patients identifies novel panel of circulating microRNA for early detection. Sci Rep. 2016;
63. Witwer KW. Circulating MicroRNA biomarker studies: Pitfalls and potential solutions. Clinical Chemistry. 2015.
64. Wang K, Yuan Y, Cho JH, McClarty S, Baxter D, Galas DJ. Comparing the MicroRNA spectrum between serum and plasma. PLoS One. 2012;
65. Li Y, Kowdley K V. Method for microRNA isolation from clinical serum samples. Anal Biochem. 2012;
66. Lai Y-H, Wang Y-S. Matrix-Assisted Laser Desorption/Ionization Mass Spectrometry: Mechanistic Studies and Methods for Improving the Structural Identification of Carbohydrates. Mass Spectrom. 2017;
67. Ho CS, Lam CWK, Chan MHM, Cheung RCK, Law LK, Lit LCW, et al. Electrospray ionisation mass spectrometry: principles and clinical applications. Clin Biochem Rev. 2003;
68. Zare RN, Fernández FM, Kimmel JR. Hadamard transform time-of-flight mass spectrometry: More signal, more of the time. Angewandte Chemie - International

Edition. 2003.

69. Kaiser NK, Quinn JP, Blakney GT, Hendrickson CL, Marshall AG. A novel 9.4 tesla FTICR mass spectrometer with improved sensitivity, mass resolution, and mass range. J Am Soc Mass Spectrom. 2011;
70. Hu Q, Noll RJ, Li H, Makarov A, Hardman M, Cooks RG. The Orbitrap: A new mass spectrometer. Journal of Mass Spectrometry. 2005.
71. Harvey DJ. Analysis of protein glycosylation by mass spectrometry. In: Analysis of Post-Translational Modifications Using Mass Spectrometry. 2016.
72. Zhou S, Hu Y, Desantos-Garcia JL, Mechref Y. Quantitation of permethylated n-glycans through multiple-reaction monitoring (MRM) LC-MS/MS. J Am Soc Mass Spectrom. 2015;
73. Vreeker GCM, Wuhrer M. Reversed-phase separation methods for glycan analysis. Analytical and Bioanalytical Chemistry. 2017.
74. Nagai-Okatani C, Aoki-Kinoshita KF, Kakuda S, Nagai M, Hagiwara K, Kiyohara K, et al. LM-GlycomeAtlas Ver. 1.0: A Novel Visualization Tool for Lectin Microarray-Based Glycomic Profiles of Mouse Tissue Sections. Molecules [Internet]. 2019 Aug 15 [cited 2019 Sep 13];24(16). Available from: http://www.ncbi.nlm.nih.gov/pubmed/31443278
75. Lam SK, Ng TB. Lectins: Production and practical applications. Applied Microbiology and Biotechnology. 2011.
76. Liu B, Bian H jiao, Bao J ku. Plant lectins: Potential antineoplastic drugs from bench to clinic. Cancer Letters. 2010.
77. Lavín de Juan L, García Recio V, Jiménez López P, Girbés Juan T, Cordoba-Diaz M, Cordoba-Diaz D. Pharmaceutical applications of lectins. Journal of Drug Delivery Science and Technology. 2017.
78. Weis WI, Drickamer K. Structural Basis of Lectin-Carb Ohydrate. Anna Rev Biochem. 1996;
79. Kishore U, Eggleton P, Reid KBM. Modular organization of carbohydrate recognition domains in animal lectins. Matrix Biol. 1997;
80. Drickamer K. Making a fitting choice: Common aspects of sugar-binding sites in plant and animal lectins. Structure. 1997.
81. Hashim OH, Jayapalan JJ, Lee CS. Lectins: An effective tool for screening of potential cancer biomarkers. PeerJ. 2017;
82. Van Damme EJM, Lannoo N, Peumans WJ. Chapter 3 Plant Lectins. In: Advances in

Botanical Research. Academic Press Inc.; 2008. p. 107–209.

83. Choi JW, Moon BI, Lee JW, Kim HJ, Jin Y, Kim HJ. Use of CA15-3 for screening breast cancer: An antibody-lectin sandwich assay for detecting glycosylation of CA15-3 in sera. Oncol Rep. 2018;